THE ROLE OF PARKINSON'S DISEASE ASSOCIATED MUTATIONS IN EIF4G1 ON PROTEIN TRANSLATION AND NEURODEGENERATION

by
Hao Jia

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
January, 2017

Abstract

Parkinson's disease is the second most common neurodegenerative disorder. Many clues about potential causes for PD have been gained from studies of disease-associated genetic mutations identified in familial PD cases. In 2011, missense mutations (p.Arg1502His and p.Ala502Val) of *EIF4G1* gene were reported to be associated with autosomal dominant familial PD. *EIF4G1* encodes a protein named eIF4G1 (eukaryotic translation initiation factor 4, gamma 1), which is a scaffold component in a mRNA translation initiation complex that performs a key role in both cap-dependent and cap-independent translation initiation. However, following sequencing studies indicated mutations of *EIF4G1* seemed to be rare, and standard genetic approaches had not been able to confirm the mutations are causal.

To study the pathological roles of *EIF4G1* mutations *in vivo,* we generated the *Drosophila* model expressing human eIF4G1 in dopaminergic neurons by crossing newly generated *UAS* lines with a dopaminergic neuron specific *GAL4* driver line. We tried three different approaches to obtain the *UAS* lines with relatively equal and robust transgene expression. Through characterization, we found the *RH-EIF4G1* and *AV-EIF4G1* transgenic flies had early mortality, late-onset locomotion impairment, and age-related dopaminergic neurodegeneration, while the *WT-EIF4G1* flies performed the same as the non-transgenic wild-type flies. These data suggested a causal link between *EIF4G1* mutations and PD-like phenotypes in *Drosophila* model.

To understand if and how the mutations alter the function of eIF4G1 *in vitro*, we generated *EIF4G1* mutation knock-in SHSY5Y cell lines, a human neuroblastoma cell line, using CRISPR/Cas9 technology. We found global protein translation repression in mutant

cells via ^{35}S-Met/Cys metabolic labeling, and consistent results were seen in mutation bearing transgenic flies. We found cap-dependent translation was slightly decreased in RH knock-in cells, but cap-independent translation was not affected. Whereas in AV knock-in cells, both cap-dependent and cap-independent translation were prominently decreased.

Overall, we demonstrated that *EIF4G1* mutations result in PD phenotypes in transgenic flies. Protein synthesis is decreased in mutation knock-in cells. These suggest that downregulation of mRNA translation may play a role in neurodegeneration. Further studies identifying the specific mRNAs affected by *EIF4G1* mutations will lead us to potential novel therapeutic targets for Parkinson's disease.

Valina L. Dawson, Ph.D.	Faculty Sponsor / Thesis Advisor
Ted M. Dawson, M.D., Ph.D.	Faculty Sponsor / Thesis Advisor
Christopher J. Potter, Ph.D.	Thesis Reader / Committee Member
Mark Wu, M.D., Ph.D.	Committee Member
Rachel Green, Ph.D.	Committee Member

Acknowledgements

Along the journey of completing this thesis, I have received a lot of guidance, support, and understanding from the people that I would like to acknowledge in the following. First and foremost, I would like to thank my thesis advisor, Dr. Valina Dawson and Dr. Ted Dawson, for their mentorship and support. They not only trained me how to do good science, but also guided me to become a professional in general. I greatly appreciate the resources they brought to me and the whole lab, so we always had the chance to access the cutting-edge equipment and technology. They set a high professional standard for themselves, so they are and will be my lifetime role model.

Besides my advisors, I would like to thank the rest of my thesis committee members, Dr. Christopher Potter, Dr. Mark Wu, and Dr. Racheal Green for their input that I was not able to get within the lab. Their rich knowledge and insight guided me overcoming the challenges I came across during this project. The thesis meetings I had with them were the pivotal points of this journey.

I would also like to thank the lab mates whom I worked very closely with. Dr. Ian Martin taught me all the *Drosophila* experiments as well as other molecular biology techniques. He was always patient with whatever questions I had. Dr. Senthilkumar Karuppagounder was working on the same gene as I did. He and I had a lot of insightful discussions about the gene we were working on. Dr. Stephen Eacker is an expert of protein translation. He organized a “translation club” in which we, who work on protein translation, discussed the interesting papers and our own data every other week. He also spearheaded the implementation of CRISPR/Cas9 in the lab, so he gave me many suggestions when I was developing my knock-in cell line using this technique.

In addition, I would like to thank all the other members in Dawson lab and Institute of Cell Engineering. Everyone is friendly and helpful. We together created a collaborative and productive environment.

Outside of lab, I would like to thank the Department of Physiology and Cellular and Molecular Physiology program. Our program is small, but everyone knows each other, it is a warm community like a family.

Last but not least, I would like to thank my family: for the selfless and unconditional support from my parents; for the endless love and genuine understanding from my wife. Knowing they are always there for me, no matter how hard life is, I believe I can always come through.

Table of Contents

Abstract ii
Acknowledgements iv
List of Tables viii
List of Figures ix
Chapter 1 1
Introduction 1
Pathology of Parkinson's disease 1
Genetics of Parkinson's disease 3
Identification of *EIF4G1* variants as PD-associated mutations 5
Protein translation and neurological disorders 10
Potential role of eIF4G1 in neurodegeneration 14
Chapter 2 18
Development of transgenic *Drosophila* expressing the *WT/RH/AV EIF4G1* for Parkinson's disease modeling 18
Introduction 18
Results 21
Generation of a tissue-specific *Drosophila* model of PD: *EIF4G1* transgenic flies 21
Early mortality in *EIF4G1 RH, AV* transgenic flies 30
Late-onset locomotion impairment in *EIF4G1 RH, AV* flies 32
Age-related dopaminergic neurodegeneration in *EIF4G1 RH, AV* transgenic flies 34
Discussion 37
Materials and Methods 41
Generation of the constructs for *UAS-EIF4G1* transgenic lines 41
Characterization of the *EIF4G1* transgenic *Drosophila* model 43
Materials 46
Statistical Analysis. 49
Chapter 3 50
Generation of PD-associated *EIF4G1* mutations knock-in cell line for functional study of the mutations 50

Introduction 50
Result 54
sgRNA target selection and repair template ssODN design for generating knock-in cell lines 54
sgRNA construction and functional test of the knock-in system in HEK293 cells 55
ssODN repair template and knock-in cell line generation procedure optimization 63
Effect of *EIF4G1* PD-associated mutations on protein synthesis 73
Discussion 79
Materials and Methods 83
Generation of *EIF4G1 RH/AV* knock-in SH-SY5Y cell lines 83
De Novo Protein Synthesis measurement by ^{35}S-methionine/cysteine incorporation 88
Bicistronic reporter assay 89
Materials 90
Statistical Analysis 93
Chapter 4 94
Conclusions 94
Development of transgenic *Drosophila* for *EIF4G1*-linked Parkinson's disease modeling 95
Generation of *EIF4G1* mutations knock-in cell lines for mechanistic study 98
Bibliography 103
Curriculum Vitae 113

List of Tables

Table 1. Genes with pathogenic mutations linked to Parkinson's disease 4

Table 2. Protease inhibitor cocktails 47

Table 3. Information of antibodies used in Chapter 2 48

Table 4. Primer sequences in Chapter 2 49

Table 5. Primers and ssODNs used in Chapter 3 91

Table 6. Information of antibodies used in Chapter 3 93

List of Figures

Figure 1. Mutations or variants identified in the EIF4G1 gene. 9

Figure 2. Diagram showing the protein translation initiation in eukaryotic cells. 12

Figure 3. Schematic representation of the pUAST-EIF4G1 construct for embryo injection.....23

Figure 4. Representative Immunoblots of eIF4G1 expression in transgenic flies generated by P-element-mediated transformation.....23

Figure 5. Quantitative PCR of eIF4G1 mRNAs in transgenic flies 24

Figure 6. Schematic showing phiC31-mediated integration of pUASTattB vector into the attP landing sites.....27

Figure 7. Diagram of pJFRC-EIF4G1 constructs with different copy numbers of UAS ..27

Figure 8. GAL4-induced expression of eIF4G1 in S2 cells.....28

Figure 9. eIF4G1 expression in transgenic flies.....29

Figure 10. Expression of mutant eIF4G1 induced by TH-GAL4 caused early mortality .31

Figure 11. Expression of eIF4G1 mutant protein by TH-GAL4 driver caused locomotion impairment. 33

Figure 12. Expression of eIF4G1 mutant by TH-GAL4 driver did not cause DAergic neuron death at early age. 35

Figure 13. Expression of eIF4G1 mutant caused age-related loss of TH-positive DAergic neurons.....36

Figure 14. Overexpression of eIF4G1 was hardly seen in eIF4G1 expressing plasmids transient transfection of SH-SY5Y cells.....53

Figure 15. Diagram of the design of the sgRNAs and ssODNs. 57

Figure 16. Fluorescence images showing the transfection efficiency and the expression of the constructs.....58

Figure 17. Results of functional test of sgRNAs.....59

Figure 18. Results of functional test of HDR of RH sgRNA + ssODNs in HEK cells 60

Figure 19. Results of functional test of HDR of AV sgRNA + ssODNs in HEK cells.....62

Figure 20. Results of genotyping of RH single clones by RFLP 67

Figure 21. Results of genotyping of AV single clones by Sanger sequencing of PCR products.....68

Figure 22. Optimization of ssODNs repair template for RH and AV 69

Figure 23. The NHEJ inhibitor Scr7 increases the HDR efficiency.....70

Figure 24. Identify the RH knock-in clones from genotyping by RFLP 71

Figure 25. Identify the AV knock-in clones from genotyping by RFLP and confirm by Sanger sequencing. 72

Figure 26. Reduced global translation in RH and AV knock-in cell lines. 75

Figure 27. Reduced global translation in RH and AV transgenic flies. 76

Figure 28. AV mutation reduces both cap-dependent and cap-independent translation, while RH mutation slightly reduces cap-dependent translation.....77

Chapter 1

Introduction

Pathology of Parkinson's disease

Parkinson's disease (PD) was first described by James Parkinson in his classic monograph "Essay on the Shaking Palsy" in 1817 (Parkinson 1817). It is the second-most common neurodegenerative disorder with a prevalence of approximately 1% of the population at the age of 65 and rising to 5% of those at the age of 85 (Lang and Lozano 1998, Lang and Lozano 1998). The cardinal clinical motor symptoms including resting tremor (shaking occurs when the limb is not in use), bradykinesia (a slowing of physical movement), rigidity, and postural instability (Fahn 2003). In addition, non-motor symptoms including loss of sense of smell, REM sleep-behavior disorder, constipation, mood and behavioral problems, and cognitive impairment are also documented (Chaudhuri, Healy *et al.* 2006). Some non-motor symptoms such as hyposmia, sleep disorders, and constipation precede the motor symptoms by several years. Other non-motor symptoms such as cognitive impairment appear in the late stage of disease progression (Chaudhuri, Healy *et al.* 2006). The symptoms are unique to each patient, and the progression varies from person to person.

The motor symptoms of the disease are attributable to the progressive and selective loss of dopaminergic (DAergic) neurons in substantia nigra pars compacta (Braak, Del Tredici *et al.* 2003). Dopamine, the neurotransmitter mainly produced in substantial nigra (SNpc), plays a crucial role in movement control. Loss of dopamine results in abnormal

nerve-firing patterns that cause impaired movement (Obeso, Rodriguez-Oroz *et al.* 2008). While the loss of dopamine neurons accounts for the motor dysfunction, recent studies have shown that other brain systems are also damaged, including brain regions that regulate other neurotransmitters such as serotonin, norepinephrine, and acetylcholine (Barone 2010). The dysregulations of these neurotransmitters are implicated to account for the non-motor symptoms (Barone 2010).

The histopathological hallmarks of PD include the presence of intracellular protein inclusions, termed Lewy bodies (LBs), which are predominantly composed of a protein named alpha-synuclein (α-Syn) (Spillantini, Schmidt *et al.* 1997, Spillantini, Crowther *et al.* 1998, Spillantini, Crowther *et al.* 1998). The buildup of abnormal and misfolded α-Syn forms aggregates or collections, that further accumulate to form protein threads called fibrils, the building blocks of Lewy bodies (Wakabayashi, Tanji *et al.* 2007). Several recent studies suggested that α-Syn may self-propagate from cell to cell, contributing to the progression of PD (Recasens and Dehay 2014).

Currently, there is no cure for the disease, and the available treatments can only relieve the symptoms by supplementing dopamine transmission or slowing down the degradation of dopamine (Fahn 2003). Though the drugs are effective during the early stages of the disease, over time patients manifest declined responsiveness to the treatment and suffer the severe side effects of these drugs (Fahn 2003). To develop the therapeutics that could halt the progression or even cure the devastating disease, a better understanding of the pathogenicity of PD is warranted.

Genetics of Parkinson's Disease

PD had been considered as a typical sporadic disorder until the last two decades. The identification of an increasing number of single genetic loci responsible for inherited PD with Mendelian pattern has drawn more and more attention into familial PD cases (Moore, West *et al.* 2005, Martin, Dawson *et al.* 2011). Approximately 15-25% of all PD cases have a family history, and 30% of those familial cases and 3%-5% of sporadic PD cases can be attributed to single mutation of one of several specific genes. Thus far, the genetic mutations of the following five genes have been proved to link to PD pathogenesis conclusively. These are: autosomal dominant α-Syn and the leucine-rich repeat kinase 2 (*LRRK2*); autosomal recessive PD genes parkin, *DJ-1*, PTEN-induced putative kinase 1 (*PINK1*), and *ATP13A2* (Table 1) (Golbe, Di Iorio *et al.* 1990, Polymeropoulos, Lavedan *et al.* 1997, Kitada, Asakawa *et al.* 1998, Kruger, Kuhn *et al.* 1998, Abbas, Lucking *et al.* 1999, Spira, Sharpe *et al.* 2001, Bonifati, Rizzu *et al.* 2003, Singleton, Farrer *et al.* 2003, Chartier-Harlin, Kachergus *et al.* 2004, Hatano, Li *et al.* 2004, Zarranz, Alegre *et al.* 2004, Goldwurm, Di Fonzo *et al.* 2005, Kachergus, Mata *et al.* 2005, Mata, Taylor *et al.* 2005, Nichols, Pankratz *et al.* 2005).

Gene	Mutations	Inheritance/ Phenotype	Gene product
SNCA	A53T, A30P, H50Q, G51D, E46K, Triplication, Duplication	AD, Parkinsonism with common dementia	Presynaptic protein
LRRK2	G2019S, N1437H, R141C/G/H, Y1699C, I2020T	AD, Late-onset parkinsonism	Multi-domain GTPase regulated serine/threonine kinase
PRKN	more than 100 mutations	AR, Early-onset parkinsonism	E3 Ubiquitin Ligase
PINK1	more than 40 mutations	AR, Early-onset parkinsonism	PTEN-induced putative kinase 1/Mitochondrial serine/threonine-protein kinase
DJ-1	L10P, M26I, E64D, L166P	AR, Early-onset parkinsonism	Protein deglycase, chaperone
ATP13A2	G877R, L1059, F182I G504R, Duplications	AR, Early-onset parkinsonism with Kufor-Rakeb syndrome	Lysosomal cation transporting P-type ATPase

Table 1. Genes with pathogenic mutations linked to Parkinson's disease

Monogenic PD and sporadic PD share many features including selective degeneration of DAergic neurons in the SNpc and Lewy body inclusions, suggesting that there may be overlapping signaling pathways that contribute to the pathogenesis of both forms of PD (Huang, Cheung *et al.* 2004, Dawson, Ko *et al.* 2010). Studies of PD-associated genetic mutations have brought better understanding into the potential pathogenic pathways of PD. For example, A53T mutation of α-Syn preferentially promotes the formation of α-Syn oligomers, which indicates that oligomeric α-Syn may be the primary toxic species in PD pathogenesis (Conway, Lee *et al.* 2000). Autosomal recessive mutations of parkin, an E3 ubiquitin ligase, are the most common cause of early-onset PD (Dawson and Dawson 2010). The majority of the known parkin mutations impact its E3 ligase activity, screening for the parkin substrates that accumulated in the brains of PD patients identified PARIS and its function as transcriptional repressor of peroxisome proliferator-activated receptor gamma (PPARγ) coactivator-1α (PGC-1α), the master regulator of mitochondria biogenesis (Shin, Ko *et al.* 2011). The studies of the parkin-PARIS-PGC-1α pathway have linked pathogenesis of PD to mitochondrial biogenesis. So, identification of new PD-causing gene defects and the pathways affected by these genes will provide important insight into PD pathogenesis.

Identification of *EIF4G1* variants as PD-associated mutations

Mutations of the *EIF4G1* gene were reported associated with autosomal dominant familial PD in 2011 (Chartier-Harlin, Dachsel *et al.* 2011). *EIF4G1* encodes a protein named eIF4G1 (eukaryotic translation initiation factor 4, gamma 1), one of the three

paralogs of eIF4G (Yan and Rhoads 1995, Coldwell, Sack *et al.* 2012). eIF4G is a scaffold component in a mRNA translation initiation complex (Sonenberg, Morgan *et al.* 1978), performing a key role in both cap-dependent and cap-independent translation initiation (Jackson, Hellen *et al.* 2010). Together with PABP and eIF4E, eIF4G circularizes the mRNA to stabilize it, and to promote initiation (Gallie 1991). In conventional cap-dependent translation, eIF4G functions as a molecular bridge. Working with eIF3, eIF4G assists the 43S pre-initiation complex loading to the mRNA (Andreou and Klostermeier 2013, Marchione, Leibovitch *et al.* 2013). In cap-independent translation initiation, eIF4G binds with eIF4A to form a sub-complex, and thereby initiates translation by recruiting 43S complex onto internal ribosome entry site (IRES) without the cap-binding eIF4F complex (Pelletier and Sonenberg 1988, Ali, McKendrick *et al.* 2001, Svitkin, Herdy *et al.* 2005). It has been suggested that concentration of eIF4G is a key switch regulating cap-dependent and cap-independent translation (Lomakin, Hellen *et al.* 2000).

The missense mutation c.3614G>A (p.R1205H) was initially identified in a genome-wide linkage analysis of a large French family with autosomal dominant Parkinsonism, and the same mutation was also found in another seven small families (Chartier-Harlin, Dachsel *et al.* 2011). In a subsequent screening for novel PD-associated *EIF4G1* variants, additional 4 variants, c.1505C>T (p.Ala502Val), c.2056G>T (p.Gly686Cys), c.3490A>C (p.Ser1164Arg), and c.3589C>T (p.Arg1197Trp), were identified in PD cases but not in control subjects (Chartier-Harlin, Dachsel *et al.* 2011). The A502V mutation was the most frequent substitution identified in the study, and also suggested from an ancestral founder in the haplotype analysis (Chartier-Harlin, Dachsel *et*

al. 2011). The subsequent neuropathological study showed that patients carrying one or double of these *EIF4G1* variants manifested clinical features of dementia with Lewy bodies (Fujioka, Sundal *et al.* 2013).

After the first report of *EIF4G1* mutations, other groups performed screens for additional *EIF4G1* variants. The follow-up studies validated the suggested linkage between PD and *EIF4G1* mutations, however, several studies also found the initial mutants, R1205H and A502V, in their control cohorts, so the exact PD-linked mutation sites are still controversial. Tucci *et al.* identified c.1456C > T (p.Pro486Ser) in two PD individuals out of 150 African familial PD cases, but also detected A502V variant in two controls out of 3500 European-American samples (Tucci, Charlesworth *et al.* 2012). Also, Schulte *et al.* identified seven novel nonsynonymous variants in six patients with classic PD phenotype in their 376 Central European PD cases, including p.Pro16Leu, p.Pro71Ser, p.Thr318Ile, p.Val541Gly, p.Ala550Pro, p.Gly698Ala and p.Ala717Pro as well as p.Pro486Ser (Schulte, Mollenhauer *et al.* 2012). Then they assessed the frequency of novel and previously reported variants in 975 familial and sporadic PD cases from Austria, Germany, and Hungary and 1014 general controls. p.Pro16Leu, p.Thr318Ile, p.Val541Gly, and p.Gly698Ala were validated but with very rare frequency (Schulte, Mollenhauer *et al.* 2012). Unexpectedly, the original R1205H mutation was found in three controls only (Schulte, Mollenhauer *et al.* 2012). Lesage *et al.* identified four novel nonsynonymous variants out of out of 251 autosomal dominant PD cases, mostly of French origin, including p.A433V, p.E465del, p.E462delInsGK, p.P446H, particularly the p.E462delInsGK was found in 2 affected siblings with segregation (Lesage, Condroyer *et al.* 2012). Two

previously reported variants were also found in their study, but one was in 2 isolated patients and one in a control case (Lesage, Condroyer *et al.* 2012). Using whole exome sequencing of 213 Caucasian PD patients and 272 control individuals, Nuytemans *et al.* confirmed that the initially identified p.R1205H mutation segregated in all affected members of a 3-generation family, except for an 86-year-old member (Nuytemans, Bademci *et al.* 2013). In addition, they also identified eight novel nonsynonymous variants and two small deletions in isolated PD patients: p.A425V, p.A428M, p.M432V, p.P486S, p.V541G, p.A550P, p.P1229A, p.L1233P, pG466_A468del, pE525del. Among those, both p.A425V and p.A428M were identified in the same patient (Nuytemans, Bademci *et al.* 2013). p.V541G, p.P486S, and pE525del were found in patients only, and p.M432V, p.A550P, p.P1229A, p.L1233P were present in both patients and controls (Nuytemans, Bademci *et al.* 2013) (Figure 1). Other screening studies for PD-linked *EIF4G1* mutations from different ethnic groups including Greek, Italian, African, Chinese, Japanese, and Indian, did not identify the initial R1205H and A502V mutants in either patients or control groups, nor found any novel variants with strong disease association (Chen, Chen *et al.* 2013, Li, Tang *et al.* 2013, Sudhaman, Behari *et al.* 2013, Zhao, Ho *et al.* 2013, Blanckenberg, Ntsapi *et al.* 2014, Gagliardi, Annesi *et al.* 2014, Nishioka, Funayama *et al.* 2014, Kalinderi, Bostantjopoulou *et al.* 2015). Hence, studies to date indicate that *EIF4G1* mutations are very rare, representing less than 1% of worldwide PD patients (Deng, Wu *et al.* 2015), and their penetrance is incomplete, as the similar situation in *LRRK2 G2019S* mutation (Deng, Le *et al.* 2005, Trinh, Guella *et al.* 2014), the most prevalent mutation in both familial and sporadic PD cases.

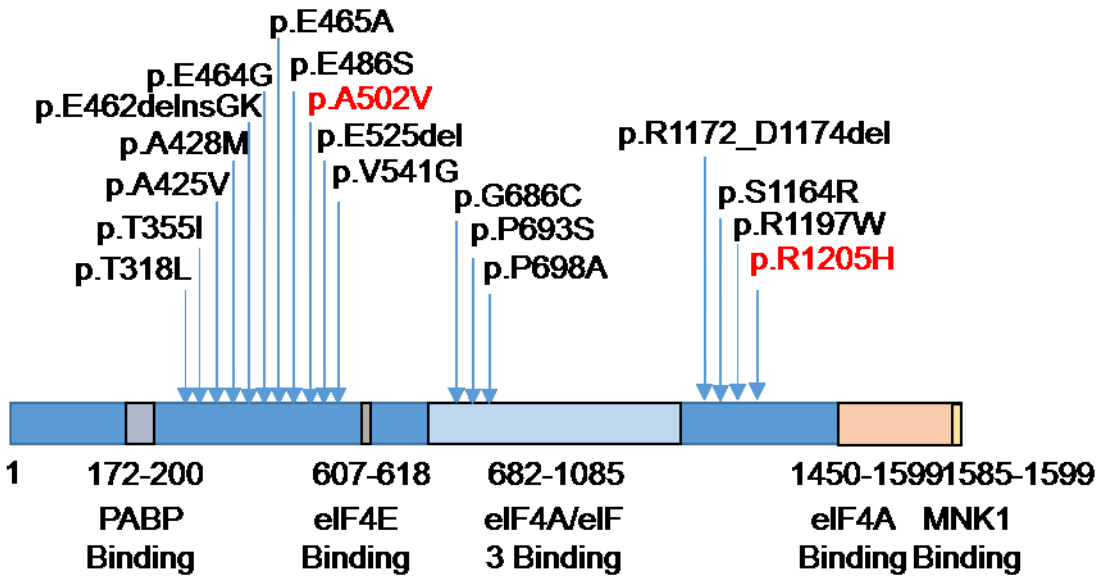


Figure 1. Mutations or variants identified in the *EIF4G1* gene.

Schematic depicting missense variants in individuals with PD and their relative location in relation to known and predicted functional domains. The amino acids number is based on NCBI Reference Sequence: NP_937884.1. The red variants are the first reported PD-associated *EIF4G1* mutations.

PABP: polyadenylate binding protein; eIF: eukaryotic translation initiation factor; MNK: Mitogen-activated protein kinase (MAPK) interacting protein kinase.

Protein translation and neurological disorders

Although the genetic studies of *EIF4G1* variants indicated the mutations are rare, we believe the study of *EIF4G1* mutations will help us know more about PD pathogenesis, as eIF4G1 is crucial in protein translation and many studies have revealed that aberrant protein translation can lead to various neurological disorders, including neurodegenerative disease.

Translational regulation is an indispensable step for gene expression during multiple life activities. Protein synthesis involves a three-stage process of initiation, elongation, and termination. Initiation is considered the rate-limiting step. Most mRNAs use the canonical 5' end-dependent scanning translation initiation pathway, while translation initiation of 5%-10% of mRNAs are cap-independent, mediated by internal ribosome entry sites (IRESs) instead (Jackson, Hellen *et al.* 2010). A brief overview of the protein translation initiation is as following (Figure 2). In eukaryotic cells, mature mRNAs are exported from the nucleus to the cytosol after processing including capping with a 7-methylguanosine residue at the 5' terminal end, polyadenylation with a poly-adenosine tail of about 200 adenylate residues at the 3'end, and splicing off the non-coding RNA introns (Izaurralde, Stepinski *et al.* 1992, Tarun and Sachs 1996). In the cytosol, at the mRNA site, forms the cap-binding eukaryotic initiation factor 4F (eIF4F) with eIF4E binding the cap, Poly-A Binding Protein (PABP) binding the Poly-A tail, eIF4A attaching and subsequently unwinding the 5' untranslated region of mRNA, and eIF4G connecting with each of the components (Izaurralde, Stepinski *et al.* 1992, Tarun and Sachs 1996, Haghighat and Sonenberg 1997). Together with the initiation factors such as eIF1, eIF1A, and eIF3, the

small ribosomal subunit (the 40S in eukaryotes) and ternary complex, comprising initiator methionyl-tRNA (Met-tRNAi-Met), eIF2 and GTP, form the 43S preinitiation complex (PIC) (Marchione, Leibovitch *et al.* 2013). The eIF4F cap complex recruits the PIC to the mRNA through the interaction of eIF4G and eIF3 and forms the 48S initiation complex (Jackson, Hellen *et al.* 2010). Once the 48S complex is loaded, it scans the mRNA from 5' to 3' until it reaches a start codon. Upon the recognition of the start codon, the GTP in the ternary complex is hydrolyzed with the aid of eIF5. Then eIF5 and eIF5B facilitate the dissociation of the initiation factors and the joining of the large ribosomal subunit (the 60S in eukaryotes). The large subunit joins to make an 80S ribosome, which has peptidyl transferase activity and thereby synthesizes polypeptide by forming peptide bonds between amino acids.

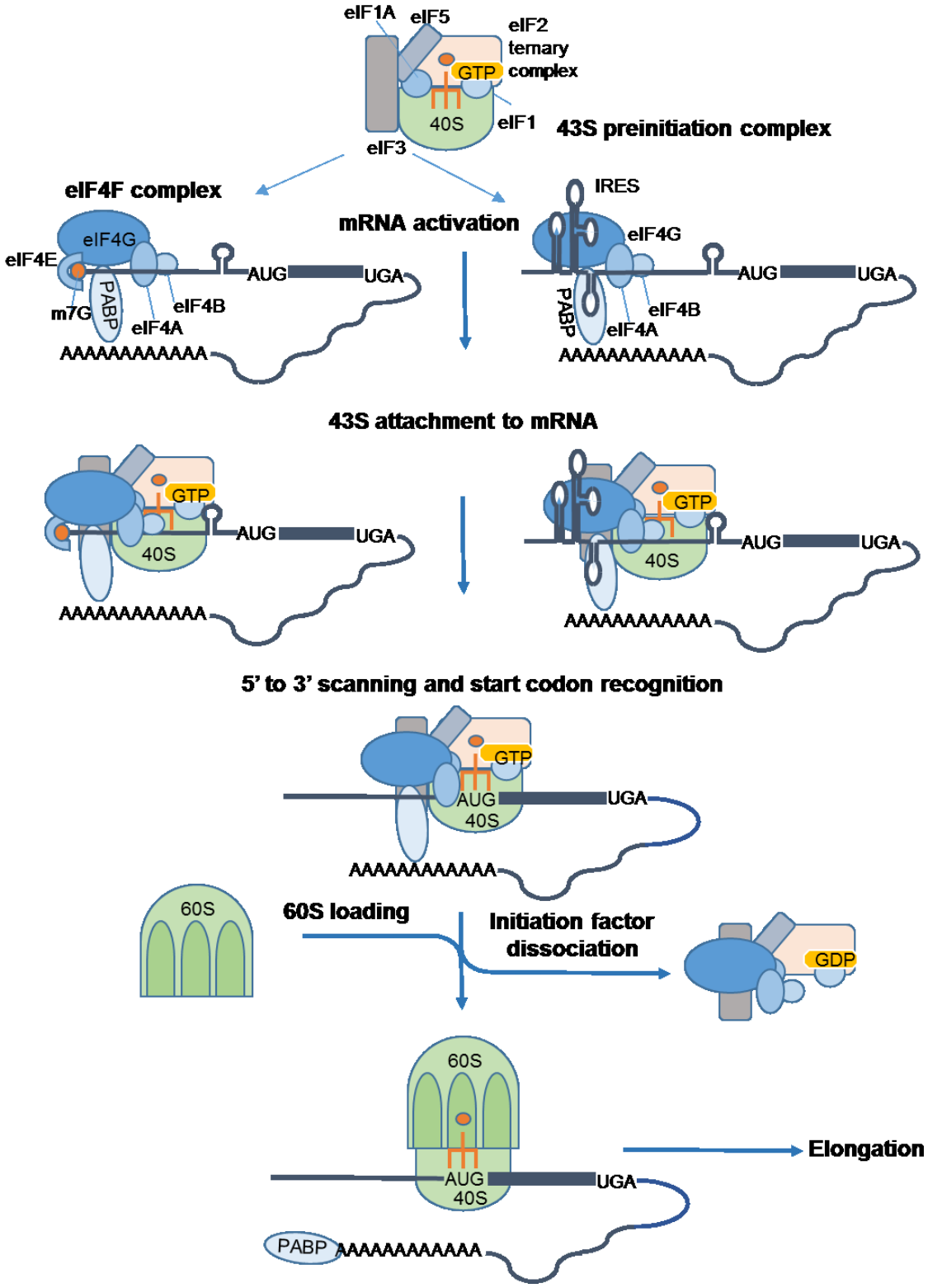


Figure 2. Diagram showing the protein translation initiation in eukaryotic cells.

There are various studies showing a causal link between neurological disorder phenotypes with dysfunction of one or more of protein translation machinery components. As described above, the level of eIF4E is crucial for cap-dependent translation, and it can be sequestered by the eIF4E-binding protein (4E-BP) from binding the eIF4F complex. It was reported that eIF4E overexpression or eukaryotic initiation factor 4E-binding protein (4E-BP) knockout mice showed the phenotypes similar to autism-spectrum disorders (ASD) (Gkogkas, Khoutorsky *et al.* 2013, Santini, Huynh *et al.* 2013). Translation of neuroligins, postsynaptic proteins that are causally linked to ASDs, was increased as eIF4E dependent. eIF2 activation is essential for the assembly of the 43S preinitiation complex and is suppressed by phosphorylation of eIF2α. eIF2α-mediated protein synthesis defects have shown to be linked with neurodegenerative diseases, such as prion-mediated neurodegeneration or Alzheimer's disease (AD) (Moreno, Radford *et al.* 2012, Ma, Trinh *et al.* 2013). eIF2α phosphorylation was elevated by the unfolded protein response (UPR) triggered by protein aggregation. Then a lot of genes that are critical for neuronal activity and synaptic function were downregulated by reduced global translation. Fragile X syndrome (FXS) and fragile X-related tremor and ataxia syndrome (FXTAS) are caused by loss of function of a single gene, fragile X mental retardation 1 (FMR1), an RNA-binding protein that has been reported functioning as a translational repressor (Sharma, Hoeffer *et al.* 2010). In the Fmr1 knock-out mouse, the mammalian target of rapamycin (mTOR) signaling pathway, which is the master regulator of translation initiation, was elevated, resulting over-activation of metabotropic glutamate receptor (mGluR) and abnormal synaptic plasticity. Recently, repeat-associated non-ATG (RAN) translation initiation has been indicated as a potential underlying mechanism to explain triplet

expansion-related homopolymeric protein disorders, such as spinocerebellar ataxia (SCA), myotonic dystrophy (DM), and amyotrophic lateral sclerosis (ALS) (Cleary and Ranum 2013).

More interestingly, recent studies about the pathogenicity of mutations of LRRK2 (Leucine-rich repeat kinase 2) have implicated the role of dysregulated protein synthesis in PD. The G2019S mutation in the kinase domain of LRRK2 is the most prevalent among both familial and sporadic PD cases (Martin, Dawson *et al.* 2011). Enhanced kinase activity of G2019 LRRK2 being responsible for its neurotoxicity has been established by multiple genetic and pharmacological studies (Smith, Pei *et al.* 2006, Lee, Shin *et al.* 2010). Identification of eukaryotic initiation factor 4E-binding protein (4E-BP) as a LRRK2 kinase substrate initially suggested the possibility of LRRK2 involvement in protein translation (Imai, Gehrke *et al.* 2008). The study, described in a recently published paper from our laboratory, provided strong evidence linking pathogenesis of PD to aberrant protein synthesis by identification of ribosomal protein s15 (rps15) as a key pathogenic kinase substrate of LRRK2 in *Drosophila* and human neuron PD models (Martin, Kim *et al.* 2014). We showed G2019S LRRK2 increased mRNA translation and protein synthesis in bulk via phosphorylation on threonine 136 of rps15, and alanine substitution rescued the neurodegeneration caused by increased kinase activity of LRRK2.

Potential role of eIF4G1 in neurodegeneration

Most of the disease cases in which eIF4G1 plays a role are cancer-related (Deng, Wu *et al.* 2015). eIF4G1 was reported overexpressed in 30% of a particular lung cancer and 80% of inflammatory breast cancer (Bauer, Diesinger *et al.* 2001, Silvera, Arju *et al.*

2009, Tu, Liu *et al.* 2010). Also, patients with higher eIF4G1 level were found associated with shorter survival time in nasopharyngeal carcinoma. PD and cancer are seemingly unrelated, but it was raised decades ago that PD patients are protected from certain type of cancers, and many epidemiological studies suggest a negative correlation between PD and cancer (Strongosky, Farrer *et al.* 2008). This link between PD and cancer indicates the hypothesis that genetic background that can protect an individual from cancer may predispose the one for PD (West, Dawson *et al.* 2005). Several genes later identified as PD-associated were studied in cancer research at the beginning. For instance, parkin and PINK1 seem to be tumor suppressor genes, while DJ-1 was initially isolated in the screen for oncogene (West, Dawson *et al.* 2005).

There are several hypotheses regarding the potential involvement of eIF4G1 in PD. Some type of PD was reported being induced by viral infection, particularly by influenza. And eIF4G1 was showed being targeted by proteases encoded by picornavirus (Foeger, Kuehnel *et al.* 2005, Jang, Boltz *et al.* 2009). The evidence that eIF4G1 is directly involved in virus-induced PD is still warranted. Several seminal studies have suggested the roles of microRNAs in PD progressing, including those showed specific microRNA targeting PD-related genes and PD-related genes regulating microRNA processing machinery (Heman-Ackah, Hallegger *et al.* 2013). And eIF4G1 has been suggested facilitating microRNA-mediated gene down-regulation. So the role of eI4G1 in PD may involve microRNA-mediated gene silencing.

As for the role of *EIF4G1* mutations in PD, while standard genetic approaches have not yet demonstrated that the mutations are causal to PD, biochemical and cellular studies

have suggested a potential role of *EIF4G1* mutations in neurodegeneration. In co-immunoprecipitation studies of eIF4G1 mutants from transfected cell lysates, the R1205H allele showed reduced interaction between eIF4G1 and eIF3E, which may hamper eIF4G1's function as a molecular bridge for 43S complex loading (Chartier-Harlin, Dachsel *et al.* 2011, Villa, Do *et al.* 2013). Furthermore, A502V eIF4G1 has been shown to perturb its binding to eIF4E, suggesting that cap-dependent translation may be affected (Sonenberg and Hinnebusch 2009, Chartier-Harlin, Dachsel *et al.* 2011). In addition, the mitochondrial membrane potential of mutant eIF4G1 overexpressed cells were substantially reduced with hydrogen peroxide treatment compared to wildtype eIF4G1 overexpressed cells, suggesting that eIF4G1 mutants may have defective stress response pathways (Chartier-Harlin, Dachsel *et al.* 2011). There is also evidence showing that eIF4G1 interacts with α-Syn (Dhungel, Eleuteri *et al.* 2015). In a yeast α-Syn model, overexpression of yeast eIF4G1 homolog TIF4631 or human eIF4G1 suppressed α-Syn toxicity, while upregulation of R1205H eIF4G1 was not able to suppress the toxicity.

Also, cellular level of wildtype eIF4G1 was reported to be crucial in its neuroprotective role in the ischemia-induced neuronal death. Vosler *et al.* showed that calpain, a protease activated by ischemia, degrades eIF4G1 and leads to persistent protein synthesis inhibition (Vosler, Gao *et al.* 2011). The authors showed that the protein synthesis rate was correlated with the viability of oxygen-glucose deprivation (OGD) treated primary neurons, and eIF4G1 overexpression increased neuronal viability after OGD treatment. These studies suggest that mutations of eIF4G1 may alter its function under stress conditions, resulting in neurodegeneration.

Further studies on *EIF4G1* mutations including genetic association to PD as well as a molecular pathologic mechanism are warranted, including examining the effects of *EIF4G*1 mutations on translation and neuronal viability in animal models. In my thesis, I demonstrated that mutations of *EIF4G1* might be disease-causing in a transgenic *Drosophila* model, and proved the eIF4G1 mutants affect protein synthesis in knock-in cells. The next chapter is about how I generated *EIF4G1* WT, RH, and AV transgenic flies and characterized their motor behavior, lifespan, and dopaminergic neuron numbers. The mutation-bearing transgenic flies showed PD-like phenotypes including locomotion defect, reduced life span, and neuronal loss compared with WT transgenic flies.

Chapter 2

Development of Transgenic *Drosophila* expressing the WT/RH/AV *EIF4G1* for Parkinson's disease modeling

Introduction

Parkinson's disease (PD) is the most common movement disorder with a prevalence of approximately 1% of the population at the age of 65 and rising to 5% of the population at the age of 85 (Lang and Lozano 1998, Lang and Lozano 1998). The cardinal symptoms including resting tremor, bradykinesia, rigidity and postural instability are attributable to the selective loss of dopaminergic neurons in substantia nigra pars compacta (Fahn 2003). PD had been considered as a typical sporadic disorder until recent two decades, identification of an increasing number of single genetic loci responsible for inherited PD with Mendelian pattern has drawn more and more attention into familial PD cases (Moore, West *et al.* 2005, Martin, Dawson *et al.* 2011). Monogenic PD and sporadic PD share many features including the pathological hallmark of PD, loss of nigrostriatal DAergic neurons. Discovery of PD-associated gene defects brings new insight into the potential pathogenic pathways of PD and other neurodegenerative diseases.

In 2011, the mutations of *EIF4G1* gene were identified as a potential cause of autosomal dominant familial Parkinson's disease in a genome-wide linkage analysis (Chartier-Harlin, Dachsel *et al.* 2011). *EIF4G1* encodes a protein named eIF4G1 (eukaryotic translation initiation factor 4, gamma 1), one of the three paralogs of eIF4G (Yan and Rhoads 1995, Coldwell, Sack *et al.* 2012). eIF4G is a scaffold component in mRNA translation initiation complex, performing a key role in both cap-dependent and

cap-independent translation initiation (Sonenberg, Morgan *et al.* 1978, Jackson, Hellen *et al.* 2010). In cap-dependent translation, together with PABP and eIF4E, eIF4G circularizes the mRNA to stabilize it, and to promote initiation (Gallie 1991). Working with eIF3, eIF4G assists the 43S pre-initiation complex loading to the mRNA (Andreou and Klostermeier 2013, Marchione, Leibovitch *et al.* 2013). In cap-independent translation initiation, eIF4G binds with eIF4A to form a sub-complex, thereby initiates translation by recruiting 43S complex onto internal ribosome entry site (IRES) without the cap-binding eIF4F complex (Pelletier and Sonenberg 1988, Ali, McKendrick *et al.* 2001, Svitkin, Herdy *et al.* 2005). However, following sequencing studies indicated mutations of *EIF4G1* seemed to be rare, standard genetic approaches have not been able to confirm the mutations are causal (Deng, Wu *et al.* 2015). The question we wanted to ask was if the mutations could cause PD-like phenotypes in genetically engineered animal models.

Currently, there are no ideal murine models for PD, so the development of *Drosophila* model fills some of the gaps as a powerful tool to study PD. Flies propagate fast, and their life spans are within 2-3 months, this is a huge time advantage for the study of age-related disease compared with around 24 months of lifespans of mice (Vanhauwaert and Verstreken 2015). Flies are capable of performing complex motor behaviors like climbing and flying, and their brains are complex enough to make these behaviors relevant to humans (Ali, Escala *et al.* 2011). Their DAergic neurons are clustered and amenable to staining and imaging, by manually counting the DAergic neurons one can easily evaluate if there is DAergic neuron loss which is the hallmark of Parkinson disease phenotype (Marsh and Thompson 2006). And the availability of very potent genetic tools including

mutagenesis by transposons, loss-of-function screen by RNAi, and *GAL4* and *UAS* system (Venken and Bellen 2007).

One of the most widely and powerful uses of the *GAL4*/*UAS* system is to generate a transgenic animal model for tissue-specific expression of a chosen gene (Elliott and Brand 2008). The yeast *GAL4* transcription activator can be induced by the temporal and spatial endogenous enhancer/promoter elements. The transgene is cloned into a construct with minimal promoter linked with upstream activator sequences (*UAS*) which can be specifically recognized by *GAL4*. *GAL4* binds the *UAS* and induces *UAS*-linked gene expression. When a *UAS*-linked transgenic line is crossed with tissue-specific promoter-linked *GAL4* driver line, the F1 progeny bearing the two transgenes will express the *UAS*-linked gene in a tissue-specific manner. The *Drosophila* research community has accumulated many different types of *GAL4* lines under different tissue specific promoter or other inducible promoters including heat shock, steroid, etc. Individual researchers just need to generate *UAS* lines by cloning their gene of interest into *UAS* construct then crossing the *UAS* lines with different *GAL4* lines depending on their research purposes. As for the *UAS* lines, the insertion location of *UAS*-linked gene into the fly genome can affect the expression of the transgene. There have been a variety of *UAS* constructs utilizing different systems to control the insertion locations.

In my thesis project, I have tried three different techniques generating the *UAS* lines, and I have selected the optimal strategy for making the *UAS*-linked human wild type (WT) *EIF4G1*, *R1205H EIF4G1*, and *A502V EIF4G1* lines and crossed with *tyrosine hydroxylase (TH)-GAL4* lines to induce transgene expression in dopaminergic neurons.

During characterization, our tissue-specific transgenic flies showed age-dependent locomotion defect, reduced lifespan, and neuronal degeneration.

Results

Generation of a tissue-specific *Drosophila* model of PD: *EIF4G1* transgenic flies

We have generated tissue-specific *WT-EIF4G1*, *R1205H-EIF4G1*, and *A502V-EIF4G1* transgenic flies using the *GAL4/UAS* system. Since the *TH-GAL4* line, in which *GAL4* is regulated under the promoter of *Drosophila tyrosine hydroxylase (DTH)* gene (Friggi-Grelin, Coulom *et al.* 2003), is already available in the Bloomington *Drosophila* Stock Center, I focused on generating the *UAS-WT-EIF4G1, UAS-R1205H-EIF4G1,* and *UAS-A502V-EIF4G1* lines. *UAS* lines are made through germline transformation by embryo injection of the construct with the gene of interest and transformation elements. The traditional and widely used transformation method relies on *P*-element transposon system (Castro and Carareto 2004). *P*-element, originally found in the fly genome, contain two terminal repeats and are capable of transposition with the DNA flanked by the terminal repeats when catalyzed by *P* transposase (Rubin and Spradling 1982). The P-element transformation vector has been developed with *P*-element ends and a range of visible or selectable marker genes, the most widely used is the *mini-white* gene (Karess and Rubin 1984, Klemenz, Weber *et al.* 1987). Individual experimenter just needs to clone the DNA of interest into the *P*-element transformation vector and co-inject with the helper plasmid encoding the *P* transposase into the embryo, then select for the marker gene for successfully transformants.

As the R1205H mutation was the first variant identified, we started with making the *WT-EIF4G1* and *R1205H-EIF4G1 UAS* line. We cloned the *WT EIF4G1* and *RH EIF4G1* into the *pUAST* vector (Figure 3) and obtained seven transformants. We crossed the *UAS* lines with the pan-neuronal *GAL4* driver line, *Embryonic Lethal Abnormal Visual (Elav)-GAL4* to obtain the progenies with transgene expression in the whole brain (Koushika, Lisbin *et al.* 1996). Then we checked the transgene expression by collecting the fly head and running the brain lysate on SDS-PAGE. Western blot by anti-human-eIF4G1 showed the expression levels cross the lines. As P-element induced transgene insertion is random, ideally, we need to have at least two lines of *WT EIF4G1* with transgene expression level similar to at least two lines of *RH EIF4G1*, but we could not find the desired lines to proceed (Figure 4). We also detected the mRNA level of the transgene by real-time PCR and found the amounts of mRNA of transgene across the lines were consistent with the protein levels (Figure 5). It indicated that the protein expression differences were caused by the transcription level, since all mRNAs of the transgene had same sequence except for the point mutation, the variation of mRNA levels may be attributed to the endogenous genomic elements as the transgene were inserted in different locations of the genome.

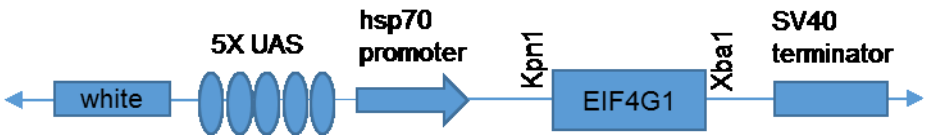


Figure 3. Schematic representation of the *pUAST-EIF4G1* construct for embryo injection.

The human *EIF4G1* CDS was subcloned in between the Kpn1 and Xba1 multi-cloning sites of the backbone p*UAS*T vector (Brand and Perrimon 1993), which contains five tandemly arrayed optimized *GAL4* binding sites (*UAS*), hsp70 promoter, and SV40 small T intron. These features are included in a *P*-element vector (pCaSpeR3) containing the P-element ends and the selective marker *white* gene.

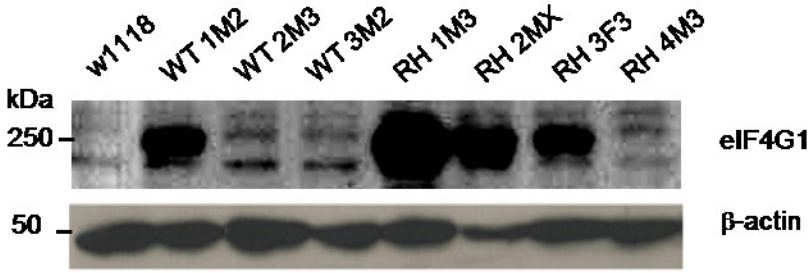


Figure 4. Representative Immunoblots of eIF4G1 expression in transgenic flies generated by P-element-mediated transformation.

Three lines of *UAS-WT-EIF4G1*, three lines of *UAS-RH-EIF4G1*, and non-transgenic *w1118* line were crossed with *ELAV (embryonic lethal abnormal visual system)-GAL4* line, with pan-neuronal expressed *GAL4*. Thirty F1 progenies bearing both *GAL4* and *UAS* were selected for the marker gene, fly heads were harvested and homogenized, protein lysates

were analyzed by SDS-PAGE, and human eIF4G1 expression levels were detected by immunoblot, β-actin level was used as loading control.

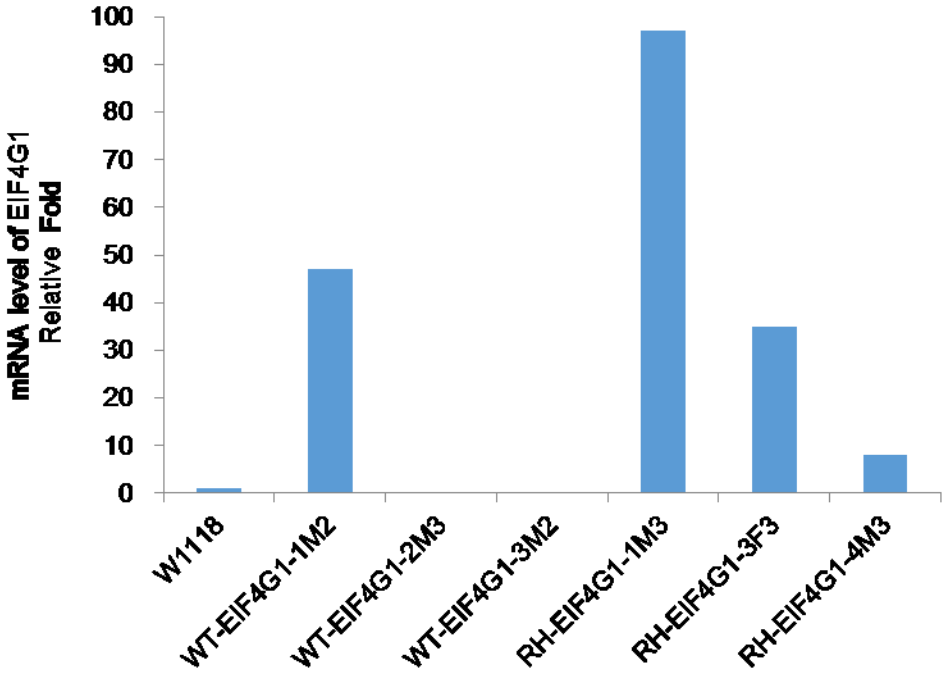


Figure 5. Quantitative PCR of eIF4G1 mRNAs in transgenic flies

For the *TH-GAL4; UAS-EIF4G1* and the control line flies, mRNAs were extracted and reverse-transcribed. cDNA was amplified by *EIF4G1* primers and tubulin (housekeeping gene control) primers and quantified. The relative mRNA levels of *EIF4G1* of all the lines were normalized to the relative mRNA level of the *w1118* line.

It has been long realized that the "position effects" of random insertion of *P*-element into the chromatin cause the various expression level of the transgene (Bellen, Levis *et al.* 2004). Researchers have developed the site-specific integration system induced by the PhiC31 integrase (Groth, Fish *et al.* 2004). Isolated from the bacteriophage PhiC31, the PhiC31 integrase could mediate the unidirectional sequence-specific recombination between two largely different DNA fragments, called *attB* and *attP* (Figure 6). The efficiency of the system has been optimized by several groups. Konrad Basler's group developed a list of lines that have *attP* landing sites inserted into the *Drosophila* major chromosomes, and also the Phic31 integrase engineered into fly genome eliminate the need to inject the integrase gene (Bischof, Maeda *et al.* 2007). One can easily clone the gene of interest into the plasmid with an *attB* site and inject the construct into the flies embryo with *attP* site and PhiC31 gene incorporated in the genome. While the PhiC31 system reduces the position effect of random insertion of the transgene, the options of a broad range of expression levels of the transgene are also limited. Rubin's group found that by increasing the copy number of *UAS* from 5 to 10, the transgene expression level can be doubled (Pfeiffer, Ngo *et al.* 2010). Then they developed a set of plasmids with *UAS* sites from 5 to 40 and found that the transgene expression level correlated with the number of the *UAS* sites. So I chose to clone the *WT-EIF4G1*, *R1205H-EIF4G1*, and *A502V-EIF4G1* into the Basler's group's *pUASTattB-5XUAS* as well as the optimized *pJFRC-10XUAS, pJFRC-20XUAS* vectors (Figure 7).

We tested the constructs in a *Drosophila* S2 cell line by co-transfection of the *UAS-EIF4G1* vector with or without *GAL4* vector. Clearly, human eIF4G1 were more robustly

expressed in the *pJFRC-EIF4G1* transfected cells than the *pUASTattB-EIF4G1* transfected cells with *GAL4* induction. (Figure 8). There was no obvious difference of expression of eIF4G1 between *pJFRC-10xUAS-EIF4G1* and *pJFRC-20xUAS-EIF4G1* transfected cells. It could be explained that the protein expression machinery is saturated in both conditions due to the excessive transfection of the plasmids.

We then tested the transgene expression in the transgenic flies by crossing the *UAS* lines with *ElAV-GAL4* line and collecting the fly heads for Western Blot. In contrast to transgenic flies overexpressing eIF4G1, *ELAV-GAL4* line and non-transgenic flies were used as a negative control that do not express the exogenous *EIF4G1* transgene in the subsequent experiments and the groups are noted as "controls". The signal of eIF4G1 protein in flies made by *pUASTattB-5xUAS* was barely seen compared with the flies made by the *pJFRC* vector (Figure 9A). The transgene expressions of the flies made by *pJFRC-20xUAS* were more robust than the flies made by *pJFRC-10xUAS* and equal across the lines (Figure 9A). As eIF4G is considered highly expressed across the tissues, we chose to use the high expression transgenic flies made by *pJFRC-20xUAS*. Since the *WT-EIF4G1, RH-EIF4G1, AV-EIF4G1* DNAs are inserted in the same location of the fly genome, we proceeded with one line out of each genotype as following: WT3, RH3, AV5 (Figure 9B).

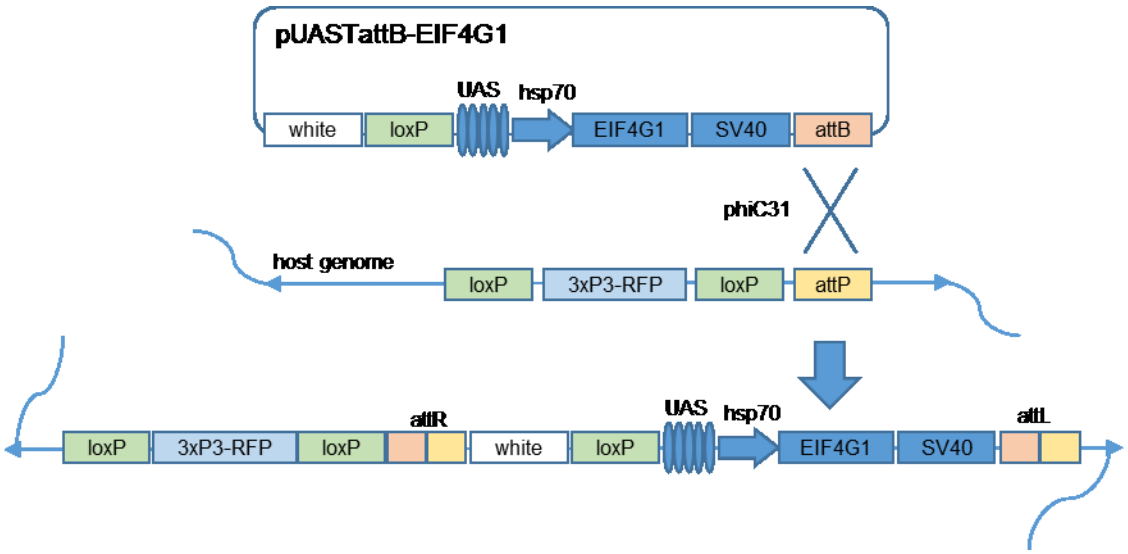


Figure 6. Schematic showing phiC31-mediated integration of *pUASTattB* vector into the attP landing sites

The *EIF4G1* CDS was sub-cloned into the *pUASTattB* vector, which contains a 285-bp *attB* fragment, a single *loxP* site in addition to the other features in the *pUAST* construct. The phiC31 integrase mediates integration of *attB*-linked plasmid into the *attP* site, creating the two hybrid attL and attR sites, which are not targetable by the phiC31 integrase. The *loxP sites* allow removing of the unnecessary sequences before or after integration of *pUASTattB*.

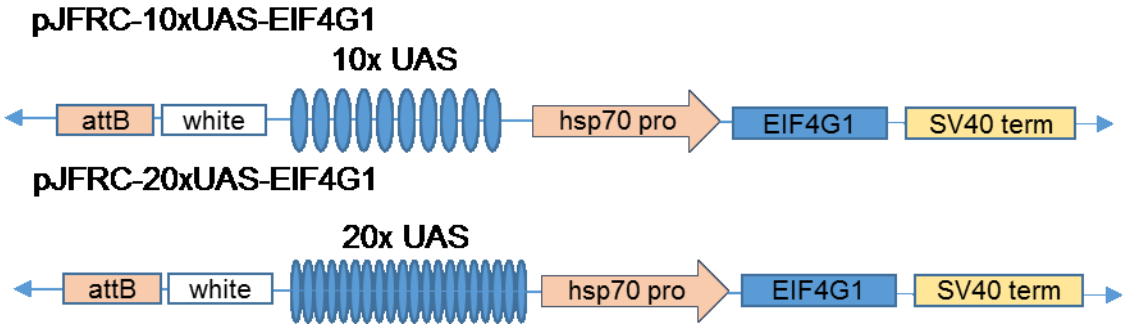


Figure 7. Diagram of pJFRC-EIF4G1 constructs with different copy numbers of *UAS*

EIF4G1 CDS fragment was sub-cloned into the *pJFRC* plasmids by the "cut-blunt-cut" method described in the Materials and Methods. *pJFRC* vectors contain different copy numbers of upstream activating sequence (*UAS*) in addition the features described in the *pUASTattB* plasmid.

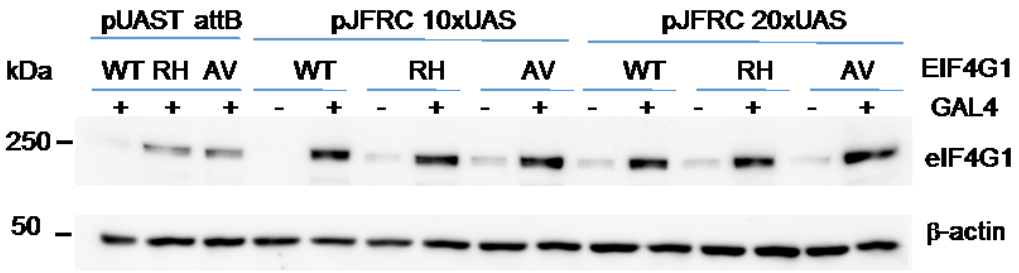


Figure 8. *GAL4*-induced expression of eIF4G1 in S2 cells

Drosophila S2 cells were transfected with *UAS-EIF4G1* constructs together with or without *GAL4* plasmid. Cells were harvested and subjected to Western blot analysis using human anti-eIF4G1 antibody.

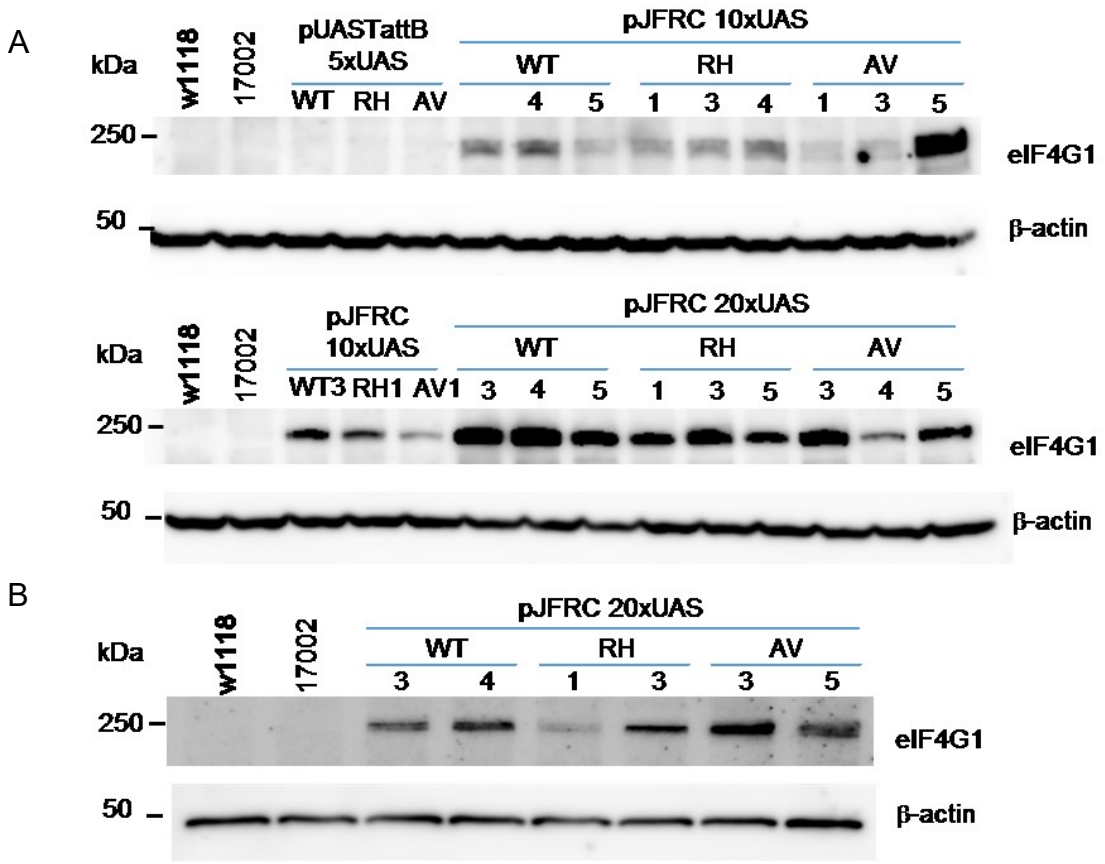


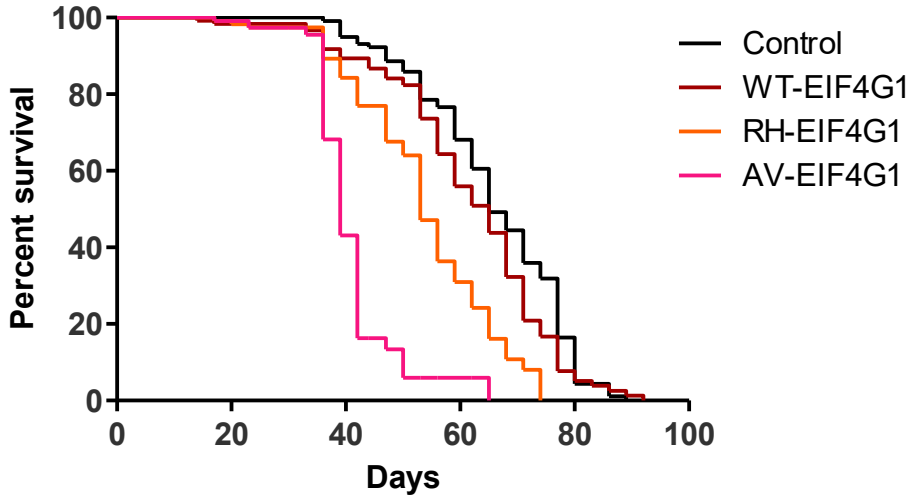
Figure 9. eIF4G1 expression in transgenic flies

(A). eIF4G1 expression in transgenic flies made by *pUASTattB, pJFRC-10xUAS,* and *pJFRC-20xUAS*. Indicated *UAS-EIF4G1* transgenic lines, wild-type w1118 line, and eIF4G knock-down 17002 line were mated with the *ELAV-GAL4* line. Fly heads extracts were fractionated and subjected to immunoblot analysis using anti-eIF4G1 antibody (20 μg of protein per lane). (B). Transgene expressions were confirmed in selected lines. WT3, RH3, and AV5 showed the equal expression of eIF4G1.

Early mortality in *EIF4G1 RH, AV* transgenic flies

To express eIF4G1 in DAergic neurons, we crossed *UAS-WT-EIF4G1, UAS-RH-EIF4G1*, and *UAS-AV-EIF4G1* with a *tyrosine hydroxylase (TH)-GAL4* line. The non-transgenic *w1118* line was also crossed with *TH-GAL4* flies as a control. About 125 progenies bearing *TH-GAL4* and *UAS-EIF4G1* or Control were collected based on the selective markers. Mortality record of all the lines were taken, and survival curve was plotted. Expression of *WT-EIF4G1* was benign for the flies as there was no significant difference between *TH-GAL4; UAS-WT-EIF4G1* flies and control flies. However, expression of either *RH-EIF4G1 or AV-EIF4G1* caused premature mortality compared with *WT-EIF4G1* expressing flies. Interestingly, *TH-GAL4; UAS-AV-EIF4G1* flies showed more severe mortality than *TH-GAL4; UAS-RH-EIF4G1* flies (Figure 10). The mean lifespans of *RH-EIF4G1* and *AV-EIF4G1* transgenic flies survived were 54 and 40 days, respectively, while the mean lifespan of *TH-GAL4; UAS-WT-EIF4G1* was 62 days. The ages at which 50% of the *RH-EIF4G1* and *AV-EIF4G1* transgenic flies survived were 54 and 39 days, respectively. Compared with 65 days of 50% survival of *TH-GAL4; UAS-WT-EIF4G1* and control lines, the mutation bearings lines had a faster mortality rate.

A



B

Genotype	Mean	Median
Control	65.8	65
WT	61.6	65
RH	53.8	53
AV	40.4	39

	Test	ChiSquare	DF	Prob>ChiSq
Control vs. WT	Log-Rank	3.12	1	0.0772
Control vs. RH	Log-Rank	44.80	1	<0.0001*
Control vs. AV	Log-Rank	166.07	1	<0.0001*
WT vs. RH	Log-Rank	25.77	1	<0.0001*
WT vs. AV	Log-Rank	154.07	1	<0.0001*
RH vs. AV	Log-Rank	78.08	1	<0.0001*

Figure 10. Expression of mutant eIF4G1 induced by TH-*GAL4* caused early mortality

(A) Survival curves of non-transgenic *w1118*, transgenic *UAS-WT-EIF4G1*, *UAS-RH-EIF4G1*, and *UAS-AV-EIF4G1*. Mortality of cohorts of 125 flies each genotype were recorded every three days. (B) Statistical analysis results comparing each line with the others.

Late-Onset locomotion impairment in *EIF4G1 RH, AV* flies

PD is typically diagnosed and characterized by motor symptoms caused by the neurodegeneration (Lang and Lozano 1998, Lang and Lozano 1998). In flies, locomotion defective phenotypes can be tested by climbing assays, as flies are negatively geotactic, i.e. flies move opposite the gravitational vector as an innate escape response (Gargano, Martin *et al.* 2005). So we performed negative geotaxis assay to test if there was behavior defect caused by expression of eIF4G1 mutations in DAergic neurons. We set up the same cross described in the survival experiment. We collected a vial of 25 flies and five vials for each line. Flies were displaced to the bottom of the vial and their climbing height within 3 seconds was recorded individually. We found that the climbing ability of all the lines declined as they were aging. And at early age, the flies expressing mutant eIF4G1 climbed as well as *WT EIF4G1* and the control flies. However, as their age increased to 5 weeks, the performance of the mutant flies declined more rapidly than WT transgenic and Control flies. Moreover, *AV-EIF4G1* flies showed more severe motor defects than *RH-EIF4G1* flies (Figure 11).

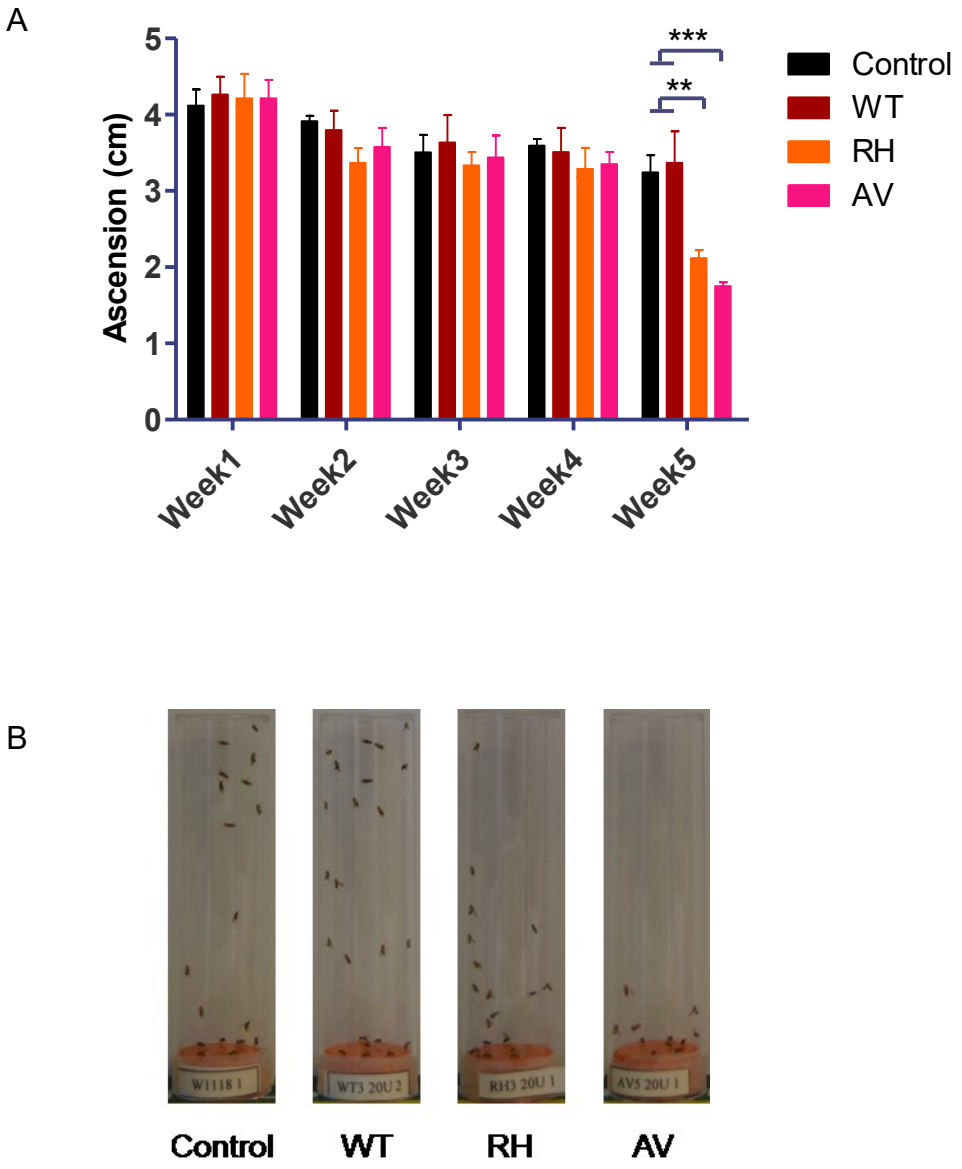


Figure 11. Expression of eIF4G1 mutant protein by TH-*GAL4* driver caused locomotion impairment.

Cohorts of 125 flies each genotype were subjected to negative geotaxis assay weekly (ANOVA, Bonferroni posttest, ***p<0.001, **p<0.01, n=5 groups of 25 flies per group). (A) Average climbing height within 3 seconds was plotted. (B) Representative pictures of each genotype at week 5.

Age-related dopaminergic neurodegeneration in *EIF4G1 RH, AV* Transgenic flies

The pathological hallmark of PD is the neurodegeneration of DAergic neurons in the SNc. To assess whether eIF4G1 mutations leads to DAergic neuron degeneration, I collected and dissected the brains of the same batch of flies as described in the negative geotaxis test at the age of 1 week and 6 week. The brains were immunostained for DAergic neuron-specific marker tyrosine hydroxylase (TH), an enzyme that catalyzes the synthesis of dopamine. I focused on the five major clusters of DAergic neurons in the central brain that can be noted according to their anatomical position and targeted by *TH-GAL4*. They are: paired posterior lateral 1 and 2 (PPL1 and PPL2); paired posterior medial 1 and 2 (PPM1/2) which are often grouped together because of their proximity; paired posterior medial 3 (PPM3); paired anterior lateral (PAL). TH-positive neurons within each cluster were manually counted through confocal Z-stacks. For one week old flies, there was no difference of DAergic neuron numbers among all the lines (Figure 12). However, at six weeks of age, the TH-positive neurons in PPL1, PPM1/2 clusters were significantly less in *RH-EIF4G1* and *AV-EIF4G1* flies than *WT-EIF4G1* and Control flies. And there was DAergic neurons loss in PPM3 of *AV-EIF4G1* flies compared with *WT-EIF4G1* and Control flies (Figure 13).

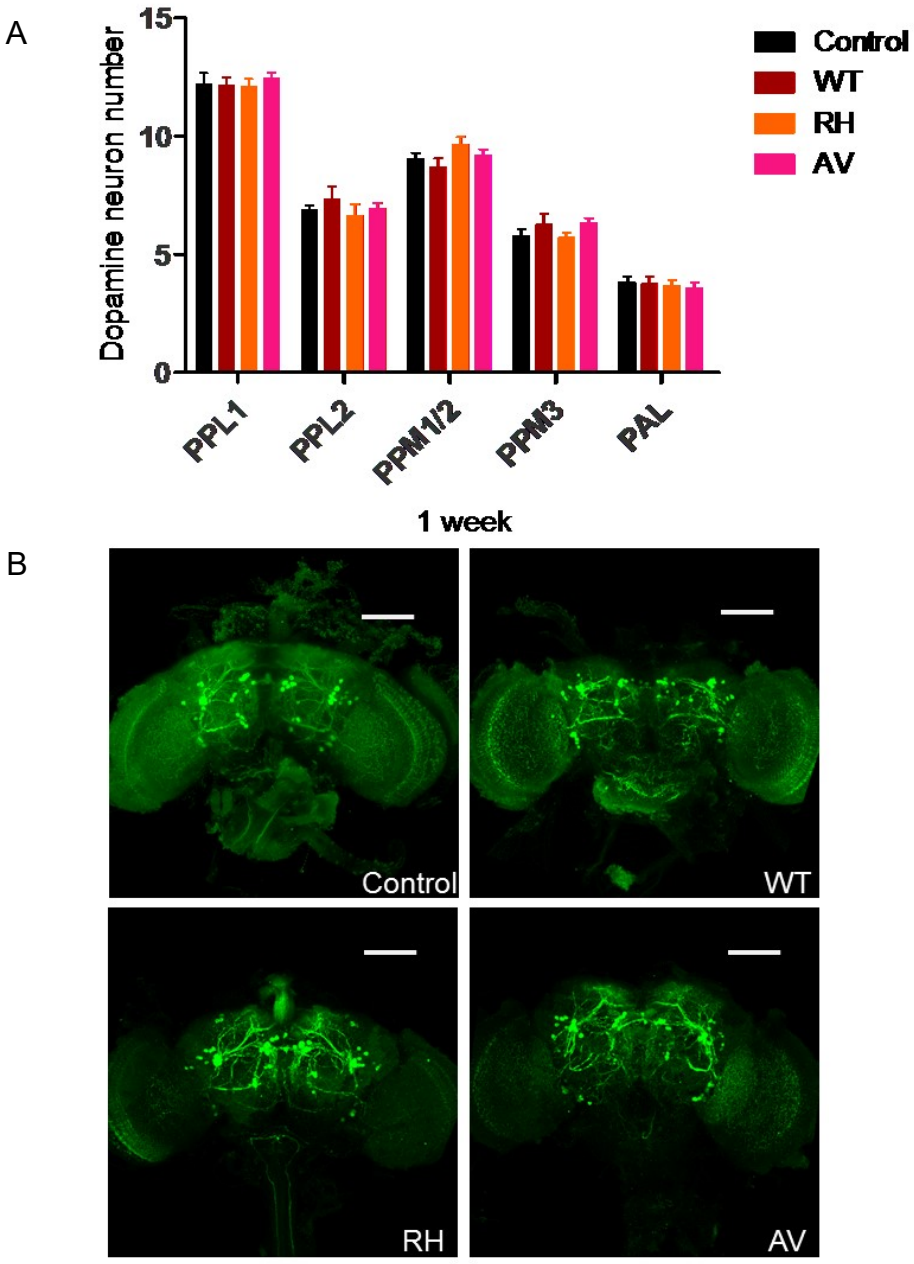


Figure 12. Expression of eIF4G1 mutant by TH-*GAL4* driver did not cause DAergic neuron death at early age.

Dissected whole brains were subjected to anti-TH immunofluorescent staining. (A) Average TH-positive neurons in each cluster of each genotype at 1-week age. (B)

Representative images of the anti-TH staining in whole brain of each genotype at 1-week age.

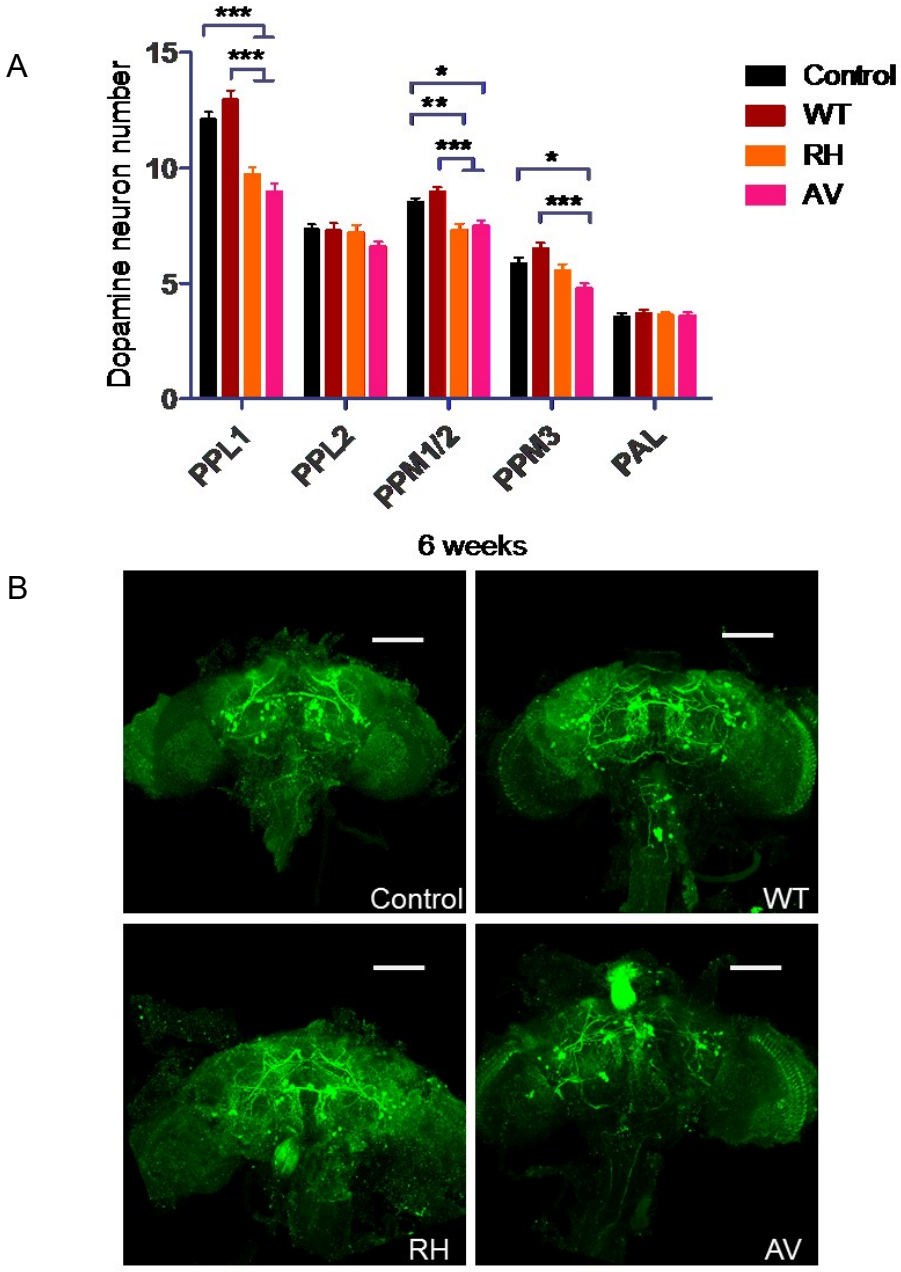


Figure 13. Expression of eIF4G1 mutant caused age-related loss of TH-positive DAergic neurons.

Dissected whole brains were subjected to anti-TH immunofluorescent staining. (A) Average TH-positive neurons in each cluster of each genotype at 5-week age. (B) Representative images of the anti-TH staining in the whole brain of each genotype at 5-week age. (ANOVA, Bonferroni posttest, ***p<0.001, **p<0.01, *p<0.05, n=10 fly brains per genotype)

Discussion

We have generated a *Drosophila* model of PD that expresses human eIF4G1 in dopaminergic neurons by crossing *UAS-EIF4G1* lines with the available *TH-GAL4* line. For the *TH-GAL4* line, yeast transcriptional activator *GAL4* gene use the enhancer/promoter of the *Drosophila Tyrosine Hydroxylase (DTH)* gene that specifically expressed in all dopaminergic cells (Friggi-Grelin, Coulom *et al.* 2003). *DTH* encodes the enzyme catalyzing the rate-limiting step in dopamine biosynthesis. To generate the construct for germline transformation for *UAS* line, we tried three different plasmids. The traditional *pUAST* plasmid utilizing *P*-element system did not give us at least two lines from each genotype showing similar transgene expression level (Figure 4). The *pUASTattB-5xUAS* plasmid was developed based on the site-specific PhiC31 integrase system. It overcomes the drawback of position effect of *P*-element, but limits the range of transgene expression level. We eventually moved to the set of plasmids developed by the Rubin lab at Janelia Research Campus, which contain different copy numbers of *UAS* sequence (Pfeiffer, Ngo *et al.* 2010). The flies generated by *pJFRC-20xUAS* showed the robust transgene expression, and the levels across the lines were equal (Figure 9).

Expression of *RH-EIF4G1* and *AV-EIF4G1* in dopaminergic neurons led to reduced life span compared with *WT-EIF4G1* and Control flies (Figure 10). Moreover *AV-EIF4G1* flies showed shorter lifespan than *RH-EIF4G1* flies. Early mortality has been observed in other *Drosophila* models expressing PD genes associated to both early-onset and late-onset parkinsonism (Lavara-Culebras and Paricio 2007, Liu, Wang *et al.* 2008). *EIF4G1 R1205H* was identified in a late-onset family, and the majority of patients' data suggested *EIF4G1 R1205H* associated with late-onset PD (Chartier-Harlin, Dachsel *et al.* 2011). These may partially reflect the smaller difference in lifespan between *RH-EIF4G1* and *WT-EIF4G1* flies. As there were few studies regarding the association of AV mutation to the age of PD onset, it is hard to draw the link between AV mutation and early onset PD just based on our fly data.

Both *RH-EIF4G1* and *AV-EIF4G1* flies had locomotion defects compared with *WT-EIF4G1* and Control flies in a negative geotaxis test (Figure 11). As the assay utilizes the natural tendency of flies to move against gravity when disturbed, it has been well accepted as the standard method to measure fly motor function. Interestingly, the motor impairment of both mutation flies manifested at a relatively late age. It indicated that the locomotion defect caused by the mutations was age-dependent. Besides negative geotaxis, we can also perform Actometer Test, which measures the mobility of flies in quiet status, to have a comprehensive profile of the locomotor activity of the flies.

We found less TH-positive neurons in both *RH-EIF4G1* and *AV-EIF4G1* flies than those in *WT-EIF4G1* and Control flies, and the loss was age-related (Figure 12). As PD is characterized by progressive degeneration of DAergic neurons among other

neurodegenerative diseases, our results indicated the *EIF4G1* mutations caused the most typical PD-like phenotype. Significant loss of dopaminergic neurons were shown in 3 clusters out of 5, and the largest percentage of neuronal loss, about 20%, was seen in PPL1 cluster. The *Drosophila* models expressing other PD-associated mutations showed DAergic neuron loss either in all or specific clusters, and percentage loss varied. It could due to either different clusters have a different tolerance for the toxicity caused by the mutations, or different clusters have different TH expression levels which cause different transgene expression level. It is known that cluster-specific neurons innervate distinct functional areas of the brain, but the particular functional effect controlled by the specific cluster is not completely characterized (Riemensperger, Isabel *et al.* 2011). We have not tested if the mutations could cause dopaminergic neuron specific loss when overexpressed in TH neurons and other brain regions. We can cross our *UAS* lines with either *dopa decarboxylase (ddc)-GAL4* driver line (Feany and Bender 2000), in which *GAL4* is expressed in both TH and serotonin (5-HT) neurons or *ElAV-GAL4* line. One caveat should be pointed out that the *GAL4* expression varies in different driver lines. We may not observe the same phenotype due to the change of expression level of the transgene. But it can be overcome by using *UAS* lines with more copy numbers of *UAS* sequence

In conclusion, we have tested a *Drosophila* model of PD-associated *EIF4G1* mutations. We provided the *in vivo* evidence of a potential causal link between *EIF4G1* mutations and PD-like phenotypes. The *RH-EIF4G1* and *AV-EIF4G1* Transgenic flies had early mortality, late-onset locomotion impairment, and age-related dopaminergic neurodegeneration, while the *WT-EIF4G1* flies performed the same as the non-transgenic

flies. While the fly model has limitations, its short lifespan and low-cost maintenance features enable us to test different hypothesis of the underlying mechanisms of pathogenicity of PD-associated mutations *in vivo*.

Materials and Methods

Generation of the constructs for *UAS-EIF4G1* Transgenic lines

Construct generation by cloning

pUAST-WT-EIF4G1: Full-length human *EIF4G1* (NM_198241.2) was cloned into the *pUAST* vector (Brand and Perrimon 1993) between Kpn1 (NEB) and Xba1 (NEB) restriction sites. *pUAST-RH-EIF4G1* and *pUAST-AV-EIF4G1* were generated by site-directed mutagenesis of the *pUAST-WT-EIF4G1* using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) and verified by sequencing.

pUASTattB-EIF4G1: Full-length human wild type or mutant *EIF4G1* were subcloned from *pUAST* constructs into the *pUASTattB* plasmids (Bischof, Maeda *et al.* 2007) between Kpn1(NEB) and Xba1 (NEB) restriction sites.

pJFRC-EIF4G1: After looking for the restriction sites on the parent vector and destination vector, we found there was only one matching site, so we used the "cut-blunt-cut" cloning method. First, the parent vector *pUASTattB-EIF4G1*and *pJFRC12-10xUAS* were cut by Sac1 (NEB) and Xho1 (NEB) respectively. Then the overhangs were made blunt by the Quick Blunting Kit (NEB). Both plasmids were purified by gel electrophoresis and proceeded with second cutting by Xba1. After the second purification, both the *EIF4G1* insert and the *pJFRC12-10xUAS* and had a blunt end and sticky end. Then standard ligation reaction was applied to generate the intact construct using the Quick Ligation Kit (NEB). The same method was applied in cloning *EIF4G1* insert into *pJFRC7-20xUAS* backbone.

Construct function test in Drosophila S2 cells

All the constructs were verified by sequencing and amplified by maxiprep (Qiagen) (primers in Table 4). Then *Drosophila* S2 cells (Gibco) were transfected with either the *pJFRC-10xUAS-EIF4G1* construct alone or together with *pCasperAUG-GAL4-X* (Vosshall, Wong *et al.* 2000) S2 cells were cultured in Schneider's Medium (Gibco) plus 10% FBS (Gibco) and 1% Penstrep (Gibco) and plated at 5*10^5 cells/well in 6-well plate (Corning). Around 1*10^6 cells in one well of six-well plate were transfected with 1ug of DNA by X-tremeGENE™ HP DNA Transfection Reagent (Roche). 24 hours after transfection cells were collected for protein level test by Immunoblot.

Immunoblot Analysis

Cells were harvested in chilled PBS and centrifuged to remove PBS. Then cells were resuspended in the cell lysis buffer and lysed on ice for 30 minutes. The cell lysate was centrifuged at 14,000 *g* for 30 min, and the supernatant was collected. Total protein amount was quantified by the BCA protein assay kit (Pierce) with BSA standards and analyzed by immunoblotting. Then the concentration of each sample was adjusted to equal by adding the cell lysis buffer. 4X Laemmli Sample buffer (Bio-Rad) containing 5% 2-Mercaptoethanol (Sigma) was added to the supernatant for sample preparation. Then the samples were boiled for 10 minutes and cooled to room temperature. The lysate samples were run in the Running Buffer through 8% SDS-PAGE gels made by 30% acrylamide/bis solution, 29:1 (3.3% C) (Bio-Rad) and transferred in the Transfer Buffer to nitrocellulose membranes. Protein bands were checked by Ponceau S dye (Sigma-Aldrich). Membranes were blocked with 5% non-fat milk in TBS-T buffer and then incubated with primary

antibodies diluted in blocking solution (2.5 % nonfat milk overnight at 4°C or 2 hours at room temperature) followed by 1 hour incubation with HRP-conjugated secondary antibodies (1:2,500). The immunoblot signal was detected using chemiluminescent substrates (Supersignal West Pico/ Femto chemiluminescent substrate, Thermo Scientific)

Characterization of the EIF4G1 Transgenic *Drosophila* model

UAS-EIF4G1 founder line generation, stock keeping, and mating set up

The selected constructs were microinjected in fly embryos (BestGene, Inc.), and we chose the 24749 strain (Basler strains (PhiC31 strains) as the *attP* host line, with *attP* inserted in 3R-86F on the genome. We received the positive marker identified and balanced adult flies. Transgene insertion and *EIF4G1* mutation were confirmed by Sanger sequencing. All the flies were maintained on standard corneal medium in 8 dram vial (Genesee Scientific) at 25 Celsius/60% relative humidity under a 12 hour light-dark cycle. Flies selected for mating were firstly transferred from vials to 6oz Square Stock Bottles (Genesee Scientific) with food and a small amount of live baker's yeast for expanding. The new generation of flies merged after 10 days. Virgin female flies were collected within 8 hours after clearing the adults at the beginning of the day. To set up the cross, around 10 virgin female flies of *UAS-EIF4G1* lines were put together with the same number of male *ELAV-GAL4* flies in the bottle. We used the estimation that 50 male flies for the experiment can be collected in one bottle of crossing. Then we set up the crossing bottles based on the number of flies we need for the experiments.

Transgene expression test

UAS-EIF4G1 lines were crossed with *ELAV-GAL4* line to generate the flies with the transgene expression in the whole brains. 30 *ELAV-GAL4; UAS-EIF4G1* flies were collected into the 1.5ml tube and froze on the dry ice. Fly heads were separated by vortexing for 30 seconds and picked manually into a new tube on the ice. Extraction buffer (1% NP-40, 50 mM Tris-HCl, 150 mM NaCl, 5 mM EGTA, protease inhibitor cocktail) was added to the tube and fly heads were homogenized by a motor-driven pestle (Kimble®). Then the tissue was lysed on ice for 30 minutes and centrifuged at 14,000 *g* for 30 min. The supernatant was collected and proceeded as the method described in the Immunoblot Analysis section. The *20xUAS EIF4G1* lines were chosen for the characterization experiments.

Real-time PCR of transgene

Fly heads were collected and homogenized, mRNAs were extracted by using the RNeasy Plus Mini Kit (Qiagen). cDNA was reversed transcribed using SuperScript® III First-Strand Synthesis Kit for RT-PCR (Thermo Scientific) using oligo(dT) primer. Transcription levels were measured by real-time PCR using SYBR green master mix (Applied Biosystems) and primers for tubulin and EIF4G1.

Survival curve

UAS-WT-EIF4G1, UAS-RH-EIF4G1, and UAS-AV-EIF4G1 flies were mated with *TH-GAL4* flies, and the progenies of *TH-GAL4; UAS-EIF4G1* had transgene expressed in their dopaminergic neurons. 5 vials of 25 flies/vial were monitored for survival. Flies were

maintained on regular food and transferred to new vials of food every three days. Mortality was scored at the same time. Kaplan-Meier survival curve was plotted and analyzed by JMP (SAS Institute, Inc., Cary, NC).

Negative geotaxis test

Cohorts of 100 flies from each genotype were maintained at 20 flies/vial and transferred to fresh food every 3 or 4 days. For each experiment, flies were transferred to the empty 8-gram vials and allowed to rest for 30 seconds. Then the vials were tapped three times so that all the flies were displaced to the bottoms of the vials to initiate climbing. After 3 seconds, the positions of the flies were captured by the fixed digital camera. The climbing assays were performed once a week using the same protocol at the same time of the day. The data analysis was done by following the method described in (Rapid iterative negative geotaxis (RING): a new method for assessing the age-related locomotor decline in *Drosophila*). The digital images were processed to 8cm high and gray sale using Photoshop (Adobe). Then the images were analyzed by using the Analyze Particle command in Scion Image (PC version of NIH Image, Scion Corporation, Frederick, MD) for the positions for each fly in the vial as X-Y coordinates. The Y values, as the indication of the climbing height within the time period, were grouped for each line at each time point. The statistics analysis was performed in Prism (GraphPad Software, San Diego, CA)

Dopamine Neuron Immunohistochemistry

Cohorts of 20 flies of each genotype were used for brain dissection for immunostaining at 1-week age and 7-week age respectively. Following the methods previously described in (Wu and Luo, 2006), fly brains were dissected into chilled 4%

paraformaldehyde, followed by fixation and permeabilization in room temperature. After 3 times of 20-minute wash in 0.3% (vol/vol) PBT (1.5 ml Triton-X 100 to 498.5 ml PB) in room temperature, 5% (vol/vol) normal goat serum (Lampire Biological Laboratories)/PBT was used in blocking. The brains were stained with primary antibody against tyrosine hydroxylase for two nights and secondary Alexa Fluor 488 goat anti-mouse IgG for another two nights. After wash, mountant reagent (ProLong® Gold Antifade Mountant, ThermoFisher) was added, and brains were mounted on the slides. Each brain was imaged by confocal microscopy. Z-stacks were taken at 1um slice intervals. TH-positive neurons were counted and recorded in each cluster of PPM1/2 (protocerebral posterior medial 1/2), PPM3, PPL1 (protocerebral posterior lateral 1), and PPL2 clusters per hemisphere.

Materials

Buffer stock solution

All stock solutions were prepared with picopure purified deionized and distilled water (ddH_2O) and filtered or autoclaved as necessary. 10X Phosphate buffered saline (PBS) was prepared by mixing 80 g of NaCl, 14.4 g of Na2HPO4, 2 g of KCl, and 2.4 g of KH2PO4 and making the final volume to 1L. The pH was adjusted to 7.4, and the solution was filtered and diluted to 1X before use. 10X electrophoresis running buffer was prepared by mixing 30.2 g of Tris base, 144.14g of Glycine, and 10 g of sodium dodecyl sulfate (SDS) per 1 L preparation. 10X transfer buffer was prepared by mixing 30.2 g of Tris base, and 144.14 g of Glycine per 1L preparation. 10X Tris Buffered Saline (TBS) buffer was

prepared by mixing 24.23 g of Tris, and 87.75 g of NaCl per 1 L preparation. For a 1X working solution of TBS-tween, 1 mL of the detergent Tween-20 was added to 1 L of diluted 1X TBS buffer. Cell lysis buffer was freshly made on the same day of each experiment, and the recipe is 50mM Tris pH 7.8, 150mM NaCl, 1%NP-40, 0.1% SDS, 5mM EGTA, and protease inhibitor cocktail. Protease inhibitor cocktails were produced by mixing various protease inhibitors as described in the table below, stored in -20 ℃ in aliquots, and freshly added to lysis buffer for total protein lysates preparation.

Chemical	Company	Concentration (1X)	Solvent
AEBSF	Sigma-Aldrich	0.5 mM	water
Aprotinin	Roche	4 ug/mL	water
Bestatin	Sigma-Aldrich	5 ug/mL	water
E-64	Roche	10 ug/mL	water
Leupeptin	Roche	10 ug/mL	water
Pepstatin	Sigma-Aldrich	1 ug/mL	ethanol

Table 2. Protease inhibitor cocktails

Antibodies

Conjugant	Antigen	Species	Supplier	Titer	Application
N/A	EIF4G1	rabbit	Cell Signaling	1:1,000	WB
N/A	Tyrosine hydroxylase	mouse	Novus Biologicals	1:1,000	IHC
HRP	α-Actin	mouse	Sigma-Aldrich	1:10,000	WB
HRP	V5	mouse	Invitrogen	1:5000	WB
HRP	FLAG	mouse	Sigma-Aldrich	1:5,000	WB
HRP	c-Myc	mouse	Invitrogen	1:5000	WB
HRP	mouse IgG	sheep	GE Healthcare	1:2,500	WB
HRP	rabbit IgG	donkey	GE Healthcare	1:2,500	WB
Alexa Fluor 488	mouse IgG	donkey	Invitrogen	1:500	IF

Table 3. Information of antibodies used in Chapter 2

HRP: horseradish peroxidase; WB: western blot; IHC: immunohistochemistry; IF: Immunofluorescence

Primers

Synthetic oligonucleotides were synthesized by the Integrated DNA Technologies. Sanger sequencing was performed at Genetic Resources Core Facility of Johns Hopkins University School of Medicine

Purpose	Sequence	Template
Verify R1205H mutation in EIF4G1	F: TCAACAAGCGGTACCCACAGAA R: CTGCACGCACTGGACTGCCTC	NM_198241.2
Verify A502V mutation in EIF4G1	F: CCAAGGAGGTGACAGCATCA R: GGCTGGATGTTCTCAGCATTG	NM_198241.2
Real-time PCR for human EIF4G1	F: TGGCTCTACCTAGCGGAACT R: TCCACCTTCTGTTCAGCGAC	NM_198241.2
Real-time PCR for *Drosophila* tubulin	F: CACTTCCAATAAAAACTCAATATG CGTGA R: ACAGTGGGTTCCAGATCCAC	NM_057424.4
Verify the 5' insertion within the backbone	GAGCGCCGGAGTATAAATAGAG	hsp70 promoter
Verify the 3' insertion within the backbone	CCATTCATCAGTTCCATAGG	SV 40 intron

Table 4. Primer sequences in Chapter 2

Statistical Analysis.

Statistical significance was determined by one-way Analysis of variance (ANOVA) and Bonferroni posttest for comparison among multiple groups of more than three, or a two-tailed nonpaired Student t-test for comparison of two groups (control and test) unless otherwise noted in the figure legends. The highest significance level was set at $p<0.05$.

Chapter 3

Generation of PD-associated *EIF4G1* mutations knock-in cell line for functional study of the mutations

Introduction

The *EIF4G1* gene, located on chromosome 3q27.1, encodes 1599 amino acids eukaryotic translation initiation factor 4G1 (eIF4G1) (Yan and Rhoads 1995). The function of eIF4G1 has been well characterized. eIF4G1 interacts with many initiation factors including eIF4E, eIF4A, eIF4B, eIF3, PABP, and some other unknown proteins serving as a scaffold protein in the initiation complex (Jackson, Hellen *et al.* 2010). The initiation complex helps the loading of the ribosome on the mRNA, which is a rate-limiting step for protein synthesis. Protein synthesis is the key process of protein homeostasis, and it is influencing critical cellular activities including metabolism, stress response, cell cycle, and cell death.

Most of the disease cases in which eIF4G1 plays a role are cancer-related (Deng, Wu *et al.* 2015). eIF4G1 was reported overexpressed in 30% of a particular lung cancer and 80% of inflammatory breast cancer (Bauer, Diesinger *et al.* 2001, Silvera, Arju *et al.* 2009). Also, patients with higher eIF4G1 level were found associated with shorter survival time in nasopharyngeal carcinoma (Tu, Liu *et al.* 2010). PD and cancer are seemingly unrelated, but it was raised decades ago that PD patients are protected from certain type of cancers, and many epidemiological studies suggest a negative correlation between PD and cancer (Strongosky, Farrer *et al.* 2008). This link between PD and cancer indicates the

hypothesis that genetic background that can protect an individual from cancer may predispose the one for PD (West, Dawson *et al.* 2005). Several genes later identified as PD-associated were studied in cancer research. For instance, *parkin* and *PINK1* seem to be tumor suppressor genes, while *DJ-1* was initially isolated in the screen for oncogene (West, Dawson *et al.* 2005). As *EIF4G1* was identified as a potential PD gene, its role in cancer may provide insight into the mechanism of the neurodegeneration caused by *EIF4G1* mutations.

The *EIF4G1* mutations, suggested as PD-associated, had never been identified before, so little was known about how the mutations could affect the functions of eIF4G1. The original paper that reported the mutations did a few biochemical studies showing that the mutations hampered the interaction between eIF4G1 and other initiation components (Chartier-Harlin, Dachsel *et al.* 2011). For instance, R1205H reduced eIF4G1 binding with eIF4E, and A502V was shown to perturb the interaction between eIF4G1 and eIF4E. Since the main functions of eIF4G1 are mRNA translation initiation related, we started to study if the eIF4G1 mutants affect protein synthesis, and how they change the translation initiation. We used the transient transfection system by transfecting SH-SY5Y cells with V5-tagged *WT-EIF4G1*, *RH-EIF4G1*, and *AV-EIF4G1*, although we found by V5 immunoblotting, expression of transfected eIF4G1 was confirmed, from eIF4G1 immunoblotting, it was hard to see the obvious overexpression of eIF4G1 (Figure 14). Transfected mutant eIF4G1 taking an unclear portion of the overall eIF4G1 pool made the results of this experiment inconsistent and highly dependent on transfection efficiency. We assumed it was because the endogenous eIF4G1 was abundantly expressed and the overall

eIF4G1 protein amount was highly regulated, it was hard to induce strong exogenous eIF4G1 expression. Those data indicated a *EIF4G1* mutation knock-in cells might be a better model. We then decided to generate the *RH-EIF4G1* and *AV-EIF4G1* knock-in cell lines by using the CRISPR/Cas system.

The CRISPR/Cas9 genomic engineering technology was developed from the RNA-based adaptive immune system in bacteria and archaea (Deveau, Garneau *et al.* 2010, Horvath and Barrangou 2010, Bhaya, Davison *et al.* 2011, Makarova, Haft *et al.* 2011). The best characterized type II CRISPR system from *Streptococcus pyogenes*, consists of the nuclease CRISPR-associated protein 9 (Cas9), the guide RNAs encoded CRISPR RNAs (crRNAs) and the required auxiliary trans-activating crRNA (tracrRNA) (Garneau, Dupuis *et al.* 2010, Sapranauskas, Gasiunas *et al.* 2011, Gasiunas, Barrangou *et al.* 2012, Jinek, Chylinski *et al.* 2012, Magadan, Dupuis *et al.* 2012). The guide RNA contains a 20-nt sequence which determines the sequence specificity (Jinek, Chylinski *et al.* 2012). Besides complementary to the guide RNA, the target must have the protospacer adjacent motif (PAM) sequence, a 3-bp NGG motif following the 20-nt sequence (Jinek, Chylinski *et al.* 2012). By optimizing the codon of Cas9 and RNA components and fusing the crRNA and tracrRNA together to create a chimeric, single-guide RNA (sgRNA), the CRIPR/Cas9 system can be implemented in various organisms (Cho, Kim *et al.* 2013, Cong, Ran *et al.* 2013, Jinek, East *et al.* 2013, Mali, Yang *et al.* 2013). Just by altering the guide RNA sequence in sgRNA, the system can be used to target theoretically any 20-nt plus NGG genomic locus to make a DNA double-stranded break (DSB). Once a DSB is made by Cas9, either one of the two major pathways for DNA damage repair will initiate: the error-prone

non-homologous end joining (NHEJ) or the high-fidelity homology-directed repair pathway (HDR) (Cho, Kim *et al.* 2013, Cong, Ran *et al.* 2013, Jinek, East *et al.* 2013, Mali, Yang *et al.* 2013). NHEJ process will result in insertion/deletion (indel) mutations which can be harnessed to mediate gene knock-out (Perez, Wang *et al.* 2008). Given different repair templates, HDR can be leveraged to make gene knock-in.

We have successfully generated the *RH-EIF4G1* and *AV-EIF4G1* knock-in SHSY5Y cell lines. And we have obtained consistent results from these cell lines. Both mutations resulted in reduced protein synthesis, and they affected both cap-dependent and cap-independent mRNA translation.

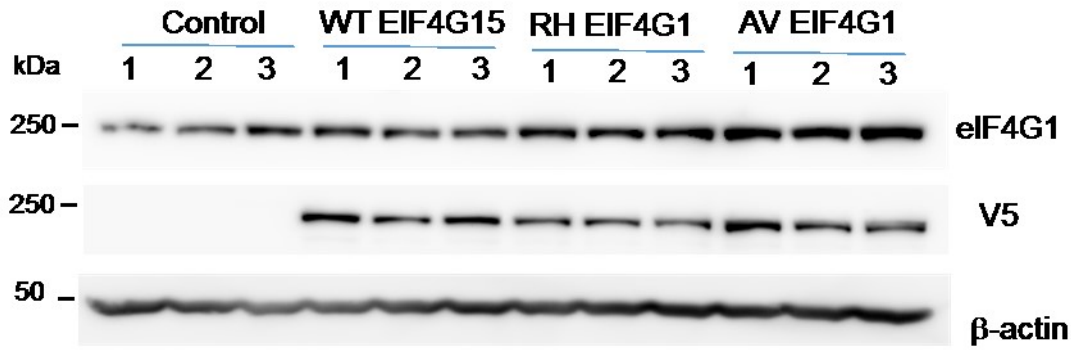


Figure 14. Overexpression of eIF4G1 was hardly seen in eIF4G1 expressing plasmids transient transfection of SH-SY5Y cells.

Representative Immunoblots showing indicated protein level. Each lane represents the cell lysate of one well of cells transfected with the indicated plasmid.

Result

sgRNA target selection and repair template ssODN design for generating knock-in cell lines

Using plasmid-based donor repair templates with longer than 500-bp homologous arms flanking the alteration sites is the traditional way to generate knock-in cells (Thomas, Folger *et al.* 1986, Hasty, Rivera-Perez *et al.* 1991). The method is good for inserting large fragments like fluorescent proteins or antibiotic resistance markers. More recently, for short modifications, using single-stranded DNA oligonucleotides (ssODNs) containing 50 ~ 100-nt homologous arm on each side flanking the alteration site has been developed as an effective and simple method (Chen, Pruett-Miller *et al.* 2011).

To make *EIF4G1* RH and AV point mutation knock-in cell lines, we chose to use the ssODNs as the repair template. For sgRNA design, we used the genomic sequence within 100-base pair proximity to the mutation site as the input. Then we used the online CRISPR design tool (http://tools.genome-engineering.org) provided by Feng Zhang's lab to generate a list of potential targeting sequences based on our input. The website provided computationally predicted off-target sites for each potential target ranked by the quantity and possibility of the predicted off-target sites (Ran, Hsu *et al.* 2013). Ideally, we want to choose the guide RNAs with as less off-target sites as possible. Based on the previous report, single-base correction rates drop as the distance between the modification site and DSB increases (Elliott, Richardson *et al.* 1998). As the Cas9 cleavage site is ~3-bp upstream of PAM, we want to select the guide RNA with targeting site locates as close as possible to the mutation site or even overlaps optimally. So we found one guide RNA with

a high off-target score of 81 (the higher the score, the less off-target sites, with 100 as the highest and 0 as the lowest, and above 50 as acceptable) and targeting site overlapping the mutation for RH knock-in. For AV knock-in, we found one sgRNA with targeting site overlapping the mutation, but its off-target score was relatively low as 64. So we added a second guide that does not overlap but has a score of 75 (Figure 15). We knew if the sgRNA targeting site on the genome remained intact after the repair template was used, a second cut could occur in the same location. So we designed a silent mutation at the PAM NGG site. Then we included 70nt homologous arm each side flanking the double stranded break for the rest of the ssODNs.

sgRNA construction and functional test of the knock-in system in HEK293 cells

For construction and delivery of CRISPR/Cas9 system with designed 20-nt guide RNA sequence, we chose one of the series of sgRNA expression constructs developed by Zhang lab, the one composed of coding sequence of Cas9, an invariant sgRNA scaffold, cloning site for specific designed sequence, 2A-GFP for selection and other expression and cloning elements, PX458. The 20-nt guide sequences were synthesized as single-stranded DNA oligos and annealed and ligated into the plasmid. The ssODNs were synthesized as 4 nmol ultramers. (Integrated DNA Technologies)

To perform functional validation of sgRNAs, we transfected HEK 293 cells with guide RNA expressing pSpCas9(BB)-2A-GFP plasmid with or without the ssODN repair template (the sense or the antisense can be considered as independent). 24 hours after transfection, we confirmed the efficiency of the transfection by checking the GFP expression under the microscope (Figure 16). 48 hours after transfection, we collected the

cells to examine if the system was functional and how the efficiency was. If the sgRNA works, it can induce a genomic inversion or deletion (indel) mutation. Indel mutations can be detected by SURVEYOR nuclease assay. We saw the desired cleavage band and estimated the efficiency ranging from 10% to 30 % (Figure 17).

To test the efficiency of the repair template, we performed restriction-fragment length polymorphism (RFLP) analysis on RH template transfected cells and Sanger sequencing on AV template transfected cells (Ran, Hsu *et al.* 2013). Basically, if the G>A mutation of R1205H was successfully made, the recognition site "TGCA" of the restriction enzyme HpyCH4V would be present. After PCR amplifying the genomic region covering the homology donor region, the DNA fragment was subject to restriction enzyme digest. By quantification of the digested DNA fragment, the efficiencies of the knock-in combination were 5% for the sense ssODN and 2% for the antisense (Figure 18). However, AV mutation does not naturally create a restriction enzyme recognition site, we performed Sanger sequencing of 25 clones of AV1 sgRNA + ssODN sense and AV2 sgRNA + ssODN antisense. Both combinations worked and estimated efficiencies were 12% and 4% respectively (Figure 19).

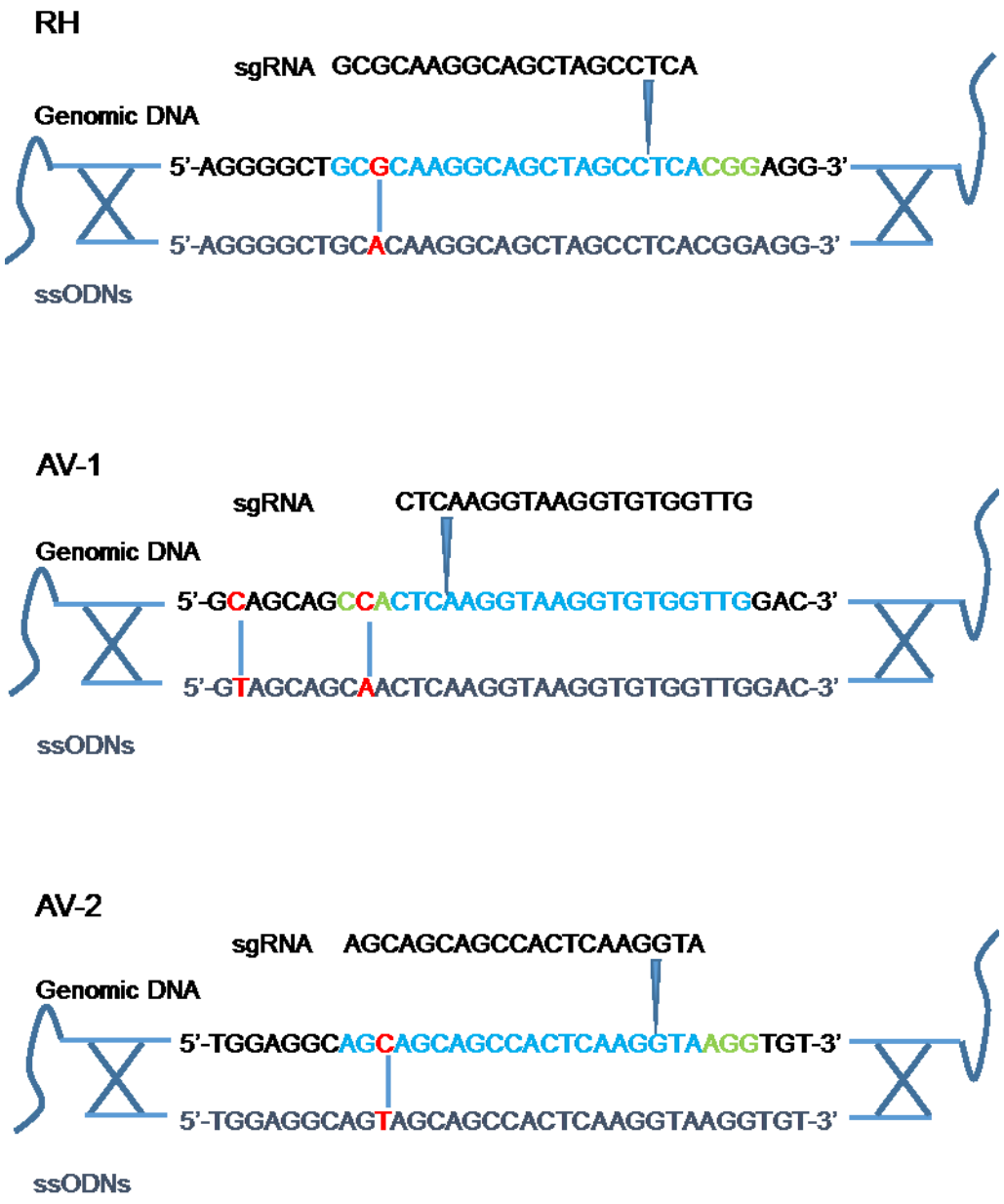


Figure 15. Diagram of the design of the sgRNAs and ssODNs.

Light blue indicates the targeting sequences of sgRNAs on the genome, green indicates the sequences of the PAMs, dark blue narrow triangle points the predicted double-stranded break site, red indicates the mutation sites.

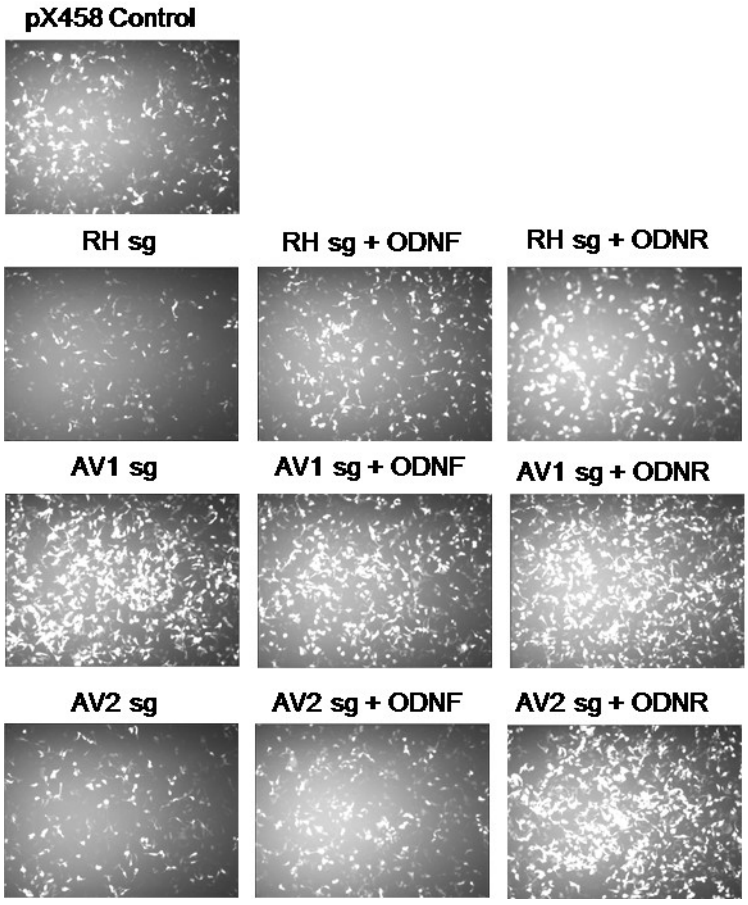


Figure 16. Fluorescence images showing the transfection efficiency and the expression of the constructs.

HEK Cells in one well of the 12-well plate was transfected with indicated single-guide RNA expressing pSpCas9(BB)-2A-GFP plasmid with or without ssODNs. ODNF: sense ODN; ODNR: anti-sense ODN.

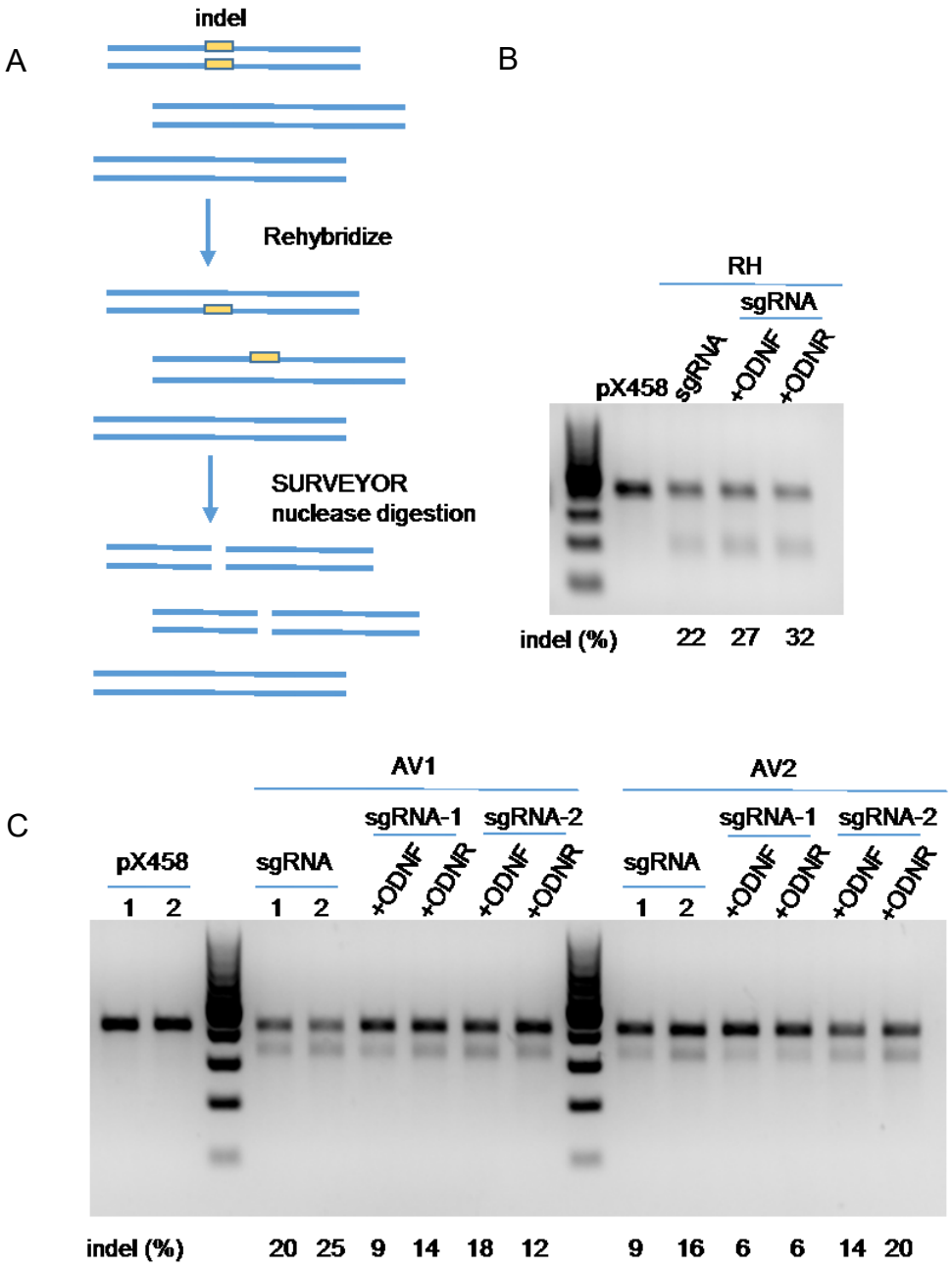


Figure 17. Results of functional test of sgRNAs

(A). Schematic of the SURVEYOR assay used to test the function of sgRNAs. (B) SURVEYOR gel shows the indel percentage of RH sgRNA +/- ssODNs treated cells. (C) SURVEYOR gels\ shows the indel percentage of AV sgRNA +/- ssODNs treated cells.

pX458 is the backbone plasmid without sgRNA. sgRNA-1 and sgRNA-2 represent the same targeting sequence but different clones.

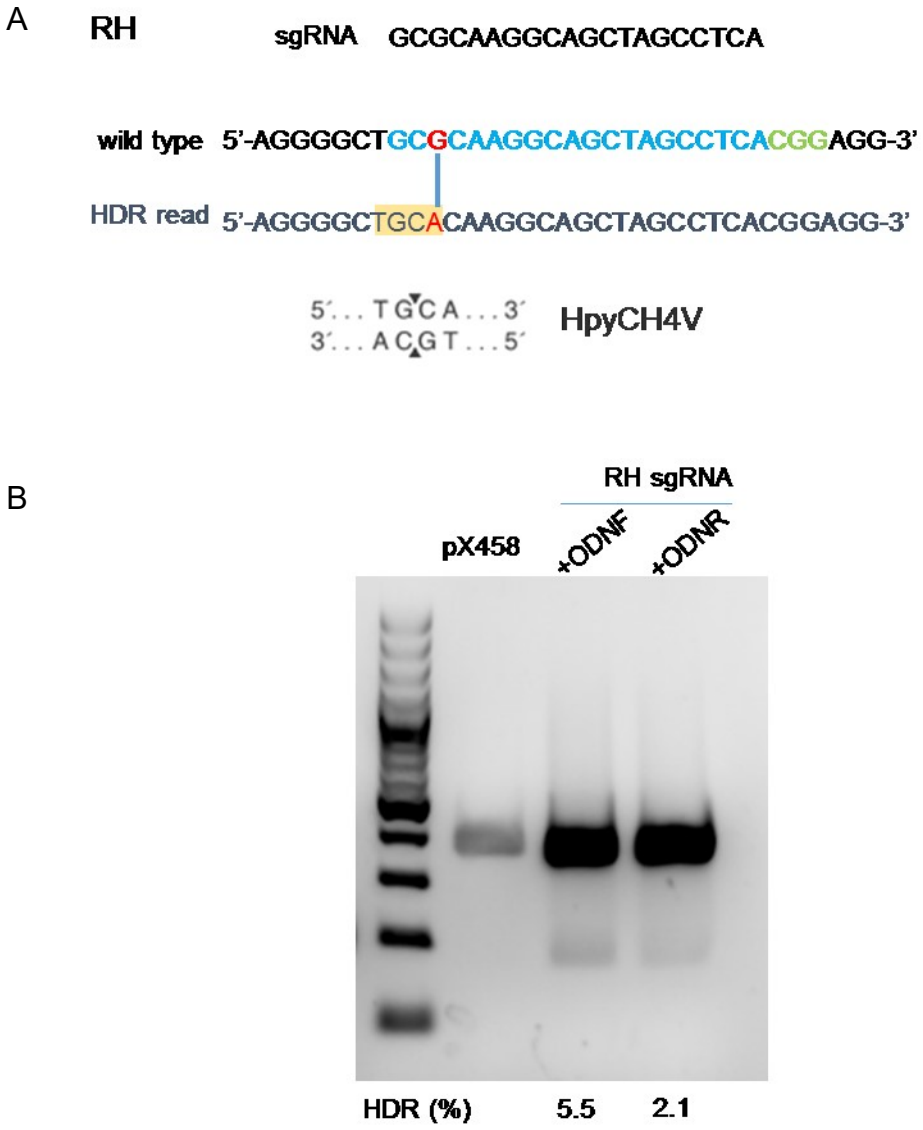


Figure 18. Results of functional test of HDR of RH sgRNA + ssODNs in HEK cells

(A) Schematic shows the designed sgRNA, genomic targeting site, and ssODN. (B) HEK cells were transfected with RH sgRNA targeting plasmid and repair template ssODNs. The restriction enzyme targeting site was introduced if the repair template was incorporated. PCR-amplified DNA fragment flanking the mutation site was digested by the indicated enzyme. Digested bands indicate the occurrence of homology-directed repair (HDR) events. ssODNs, oriented in either sense (ODNF) or anti-sense (ODNR) can be used independently to achieve efficient HDR-mediated editing.

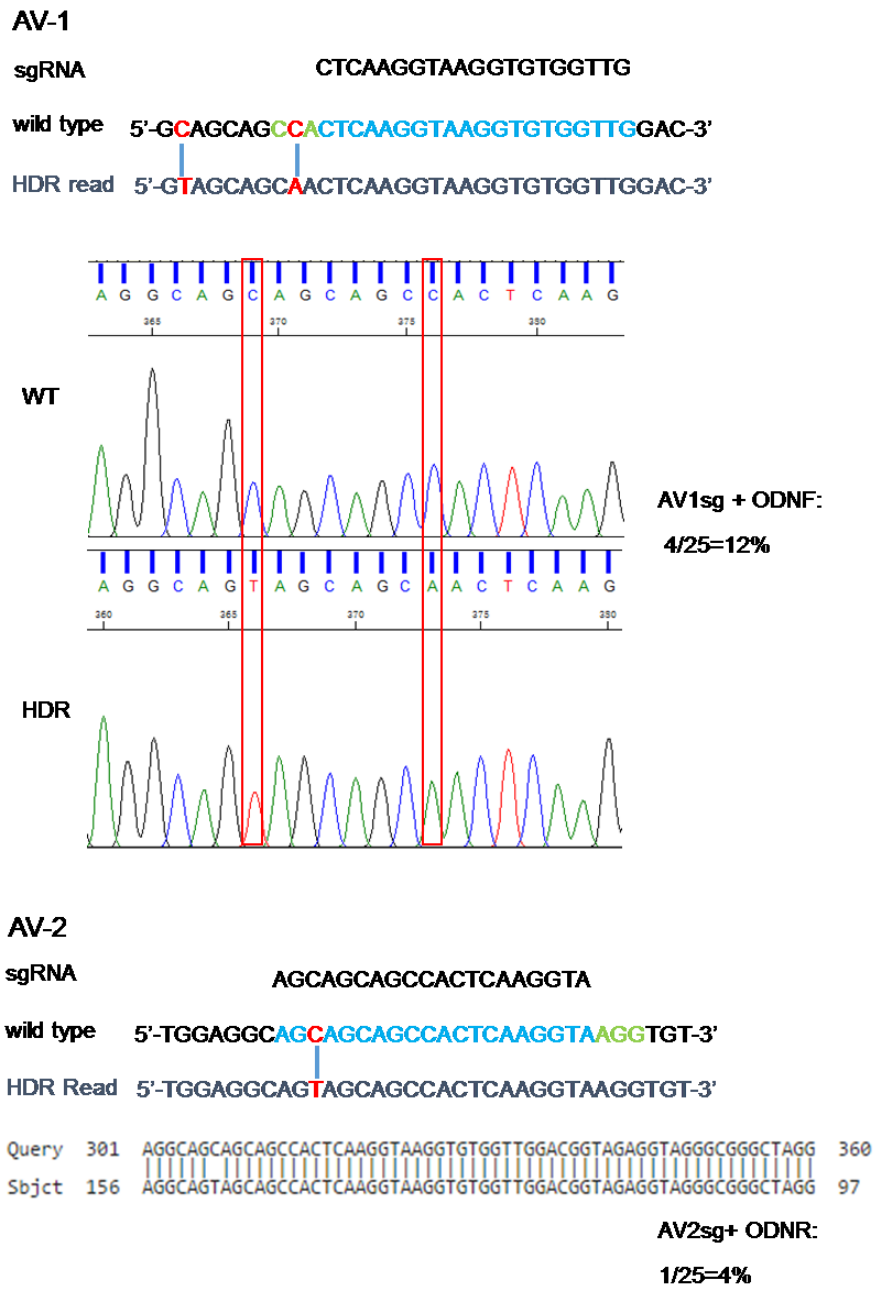


Figure 19. Results of functional test of HDR of AV sgRNA + ssODNs in HEK cells

PCR amplification of the target region on the genome from AV sgRNA + ssODNs treated cells, amplicons were TA cloned into the sequencing plasmids. 25 clones were subjected

to Sanger sequencing to assess the HDR-mediated targeting efficiency. Representative chromatogram of AV1 and sequence alignment of AV2 are shown.

ssODN repair template and knock-in cell line generation procedure optimization

Then we moved forward with the sgRNA and ssODN to transfect the human embryonic stem cells (hESCs). After the transfection, we did single cell isolation and genotyping. To our surprise, we did not identify a positive RH knock-in clone by RFLP, but when we treated the same PCR products with the restriction enzyme that could cut the wild type sequence but not any altered sequence, we found 4 out of 50 samples could not be digested, this told us the gene editing happened, but there was no intended mutation (Figure 20). From the sequencing result of AV, we did not detect any intended mutation in 300 clones treated with AV1 sgRNA + ssODN combination, but for AV2 sgRNA + ssODN, we detected the presence of the mixture peaks around our intended mutation from the chromatogram in 11 out of 200 clones. Further dissection of the sequence of the DNA samples of these 11 clones by TA cloning and plasmid sequencing, we found unwanted indel mutations in all the HDR events (Figure 21). It seemed like the HDR efficiency of AV1 sgRNA + ssODN was dramatically lower than AV2 sgRNA + ssODN in hESCs, though it was the opposite in HEK293 cells. Also, identification of unwanted indels together with intended mutation indicated that re-cut occurred at the edited site in AV2 sgRNA + ssODN treated cells.

We were aware that if the targeting sequence of sgRNA was unchanged after the integration of the repair template, a second cut could occur in the same location. For RH sgRNA and AV2 sgRNA, their binding sites was supposed to be altered after the HDR

because of the overlap between the targeting sites and the intended mutations, so it was unexpected to see indel in AV2 sgRNA + ssODN treated cells. We suspected that one nucleotide might be not sufficient to block the binding of sgRNA with the genomic DNA, so the second or more cleavages continued to happen until the NHEJ-mediated modification prevented the further cutting. So we assumed we had to induce a synonymous blocking mutation to prevent the further editing and the blocking mutation had to be sufficient enough. We knew the NGG protospacer adjacent motif (PAM) is the required element for sgRNA-mediated binding, so we decided to introduce a silent mutation in the PAM sequence of our sgRNA targeting sites. We also realized that the overall HDR efficiency was very low, and we need to use the powerful RFLP instead of Sanger sequence to genotype the AV clones. Since we could not identify a restriction enzyme site after the intended mutation was created, we had to find if we could create a restriction enzyme site close to the mutation site, at the same time the sequence remained synonymous. We found for AV1 sgRNA, it was hard to find an available synonymous restriction enzyme site between the intended mutation and PAM, and the HDR efficiency of AV1sgRNA + ssODN was suggested low. So we re-designed the repair template ssODNs for RH and AV respectively, with the intended mutation, the blocking mutation on the PAM, and the genotyping-required restriction enzyme mutation (Figure 22). And we also chose to use SH-SY5Y cells, a neuroblastoma cell line, to optimize the design and procedure.

To overcome the low efficiency problem of HDR, we also applied another improvement. It was reported that by targeting DNA ligase IV, a critical enzyme for NHEJ, using the small molecule inhibitor Scr7, the efficiency of HDR can be increased

(Maruyama, Dougan *et al.* 2015). Another group did a screening for the small molecules that boost HDR, and they found L-755,507, known as a potent and selective β3 adrenergic receptor partial agonist (Yu, Liu *et al.* 2015). We tested two drugs suggested in these studies in the RH sgRNA + ssODN transfected SH-SY5Y cells, we saw a subtle increase of the digested bands in Scr7 (1μM) and L-755,507 (1μM) treated cells (Figure 23A). After the transfection and drug treatment, we used the fluorescence sorting to select the GFP-positive cells, so only the cells with the construct transfected and expressed were enriched. We then cultured the sorted cells as a whole until they reached stable growth. At the first transfer of the cells, we did RFLP with the half of the cells to see if there were any digested bands which indicated there were knock-in cells in the population. We detected the digested band in RHsg + ssODN transfected and Scr7 treated cells, so we were confident that there were knock-in cells in this population, this time the L-755,507 treated cells showed unexpected bands, so we stayed with the Scr7 treated cells only (Figure 23B). Then we did single cell isolation by serial dilution of the cells into 384 wells. We kept monitoring the wells and identified 200 single cell developed colonies. Two weeks after we plated the cells, the cell number of single clones reached about 5000, then we collected half of the cells for genotyping and transferred the other half to larger wells to keep the clones. From the first round genotyping by RFLP, we found 3 samples out of 84 showing the small size band, indicating these three could be the knock-in clones (Figure 24). We repeated the digestion and did TA clone of the PCR products into plasmids, following by Sanger sequencing of 30 independent clones per sample to identify all alleles within the sample. We got one sample with one wild type allele and the other allele containing both the

intended mutation and the blocking mutation, and the wild type versus mutation allele ratio is approximately 1:1.

Then we repeated the same procedure with AVsgRNA + ssODN. We transfected the cells and sorted for GFP after 48 hours. Grew the cells until they were stable and used the half of the cells for population genotyping. After we confirmed there were knock-in cells in the population we did the single cell isolation and genotyping of the individual clone. This time, instead of purifying and digesting PCR products of 100 clones, we grouped PCR products of 6 samples into one sample, purified and digested the grouped samples, ran the reaction on the agarose gel. Because based our previous experience, the minimum detectable DNA amount on 2% agarose gel for low molecular size band ranges from 50 to 100 ng, if there was a positive clone out of 6 sample, as long as we loaded more than 600ng of DNA, we should be able to detect the digested band. So we identified 3 grouped samples with the positive bands and we further checked the 18 samples of individual clone from these 3 groups (Figure 25). 4 clones showed the digested bands, and we verified these 4 by cloning the PCR products into the plasmids and Sanger sequencing 30 independent bacterial clones. We found one cell line clone with one wild type allele and one edited allele containing the intended mutation, the blocking mutation, and the restriction enzyme mutation, and the allele ratio was about 1:1. So far we have successfully generated *EIF4G1 R1205H* and *EIF4G1 A502V* knock-in SHSY5Y cell lines.

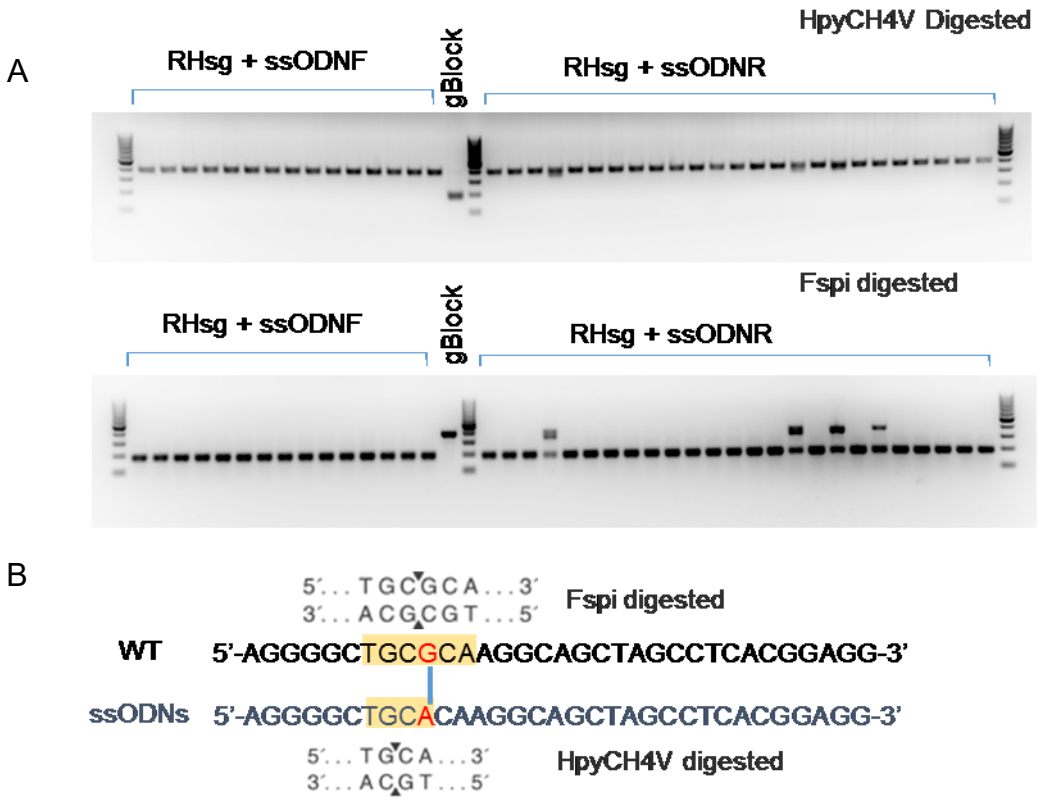


Figure 20. Results of genotyping of RH single clones by RFLP

Amplicons of DNA fragment flanking the targeting site of the genome of single clones were subjected to restriction enzyme digestion. HpyCH4V recognizes the predicted mutation sequence, Fspi recognizes the wild-type genomic sequence. gBlock is the synthesized DNA double strand fragment which has the same sequence of the PCR amplicon with the desired mutation functioning as a positive control.

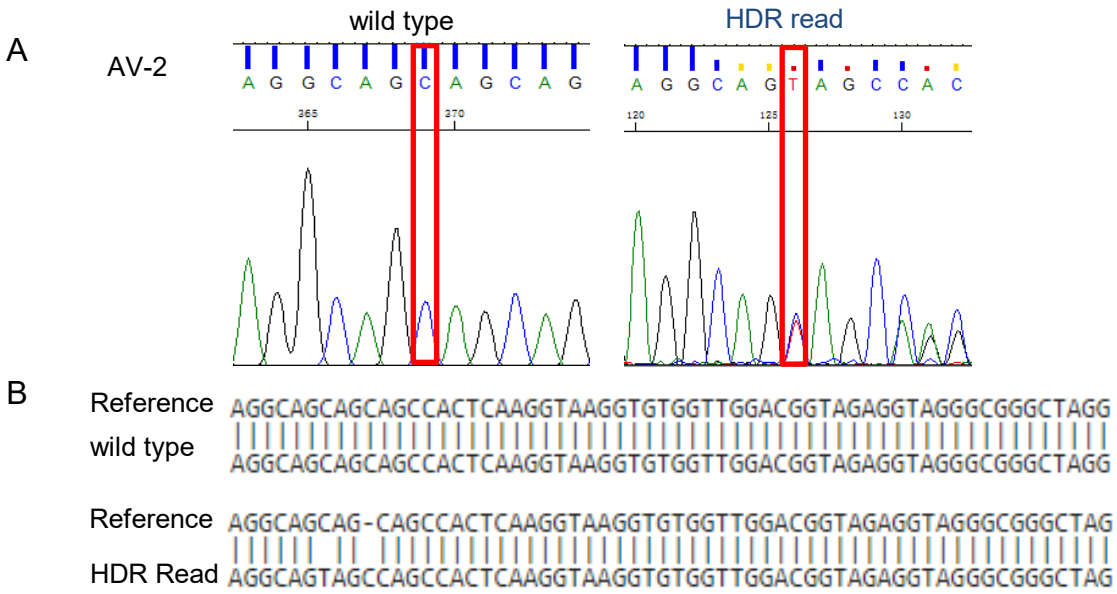


Figure 21. Results of genotyping of AV single clones by Sanger sequencing of PCR products.

(A) Representative of chromatogram of Sanger sequencing of PCR product. Amplicons of DNA fragment flanking the targeting site of genome of single clones were directly subjected to Sanger sequencing. (B) Representative sequence alignment of sequencing results of the plasmid with PCR amplicon. The clones showing mixed chromatogram was followed by TA cloning and sequencing to reveal all the sequence of all the alleles.

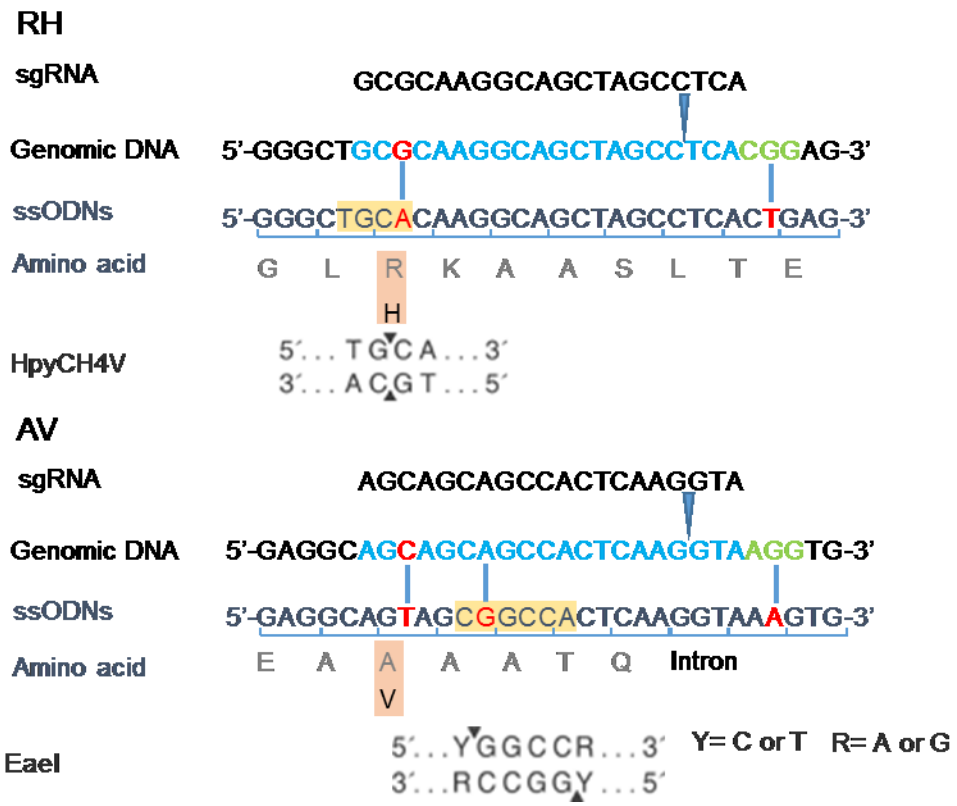


Figure 22. Optimization of ssODNs repair template for RH and AV

Diagram of the new design of the repair template ssODNs, and respective genome DNA sequences and amino acid sequences. Light blue indicates the targeting sequences of sgRNAs on the genome, green indicates the sequences of the PAMs, dark blue narrow triangle points the predicted double-stranded break site, red indicates the mutation sites, including the intended mutation, the silent blocking mutation on the PAM, and the synonymous mutation for restriction enzyme digestion.

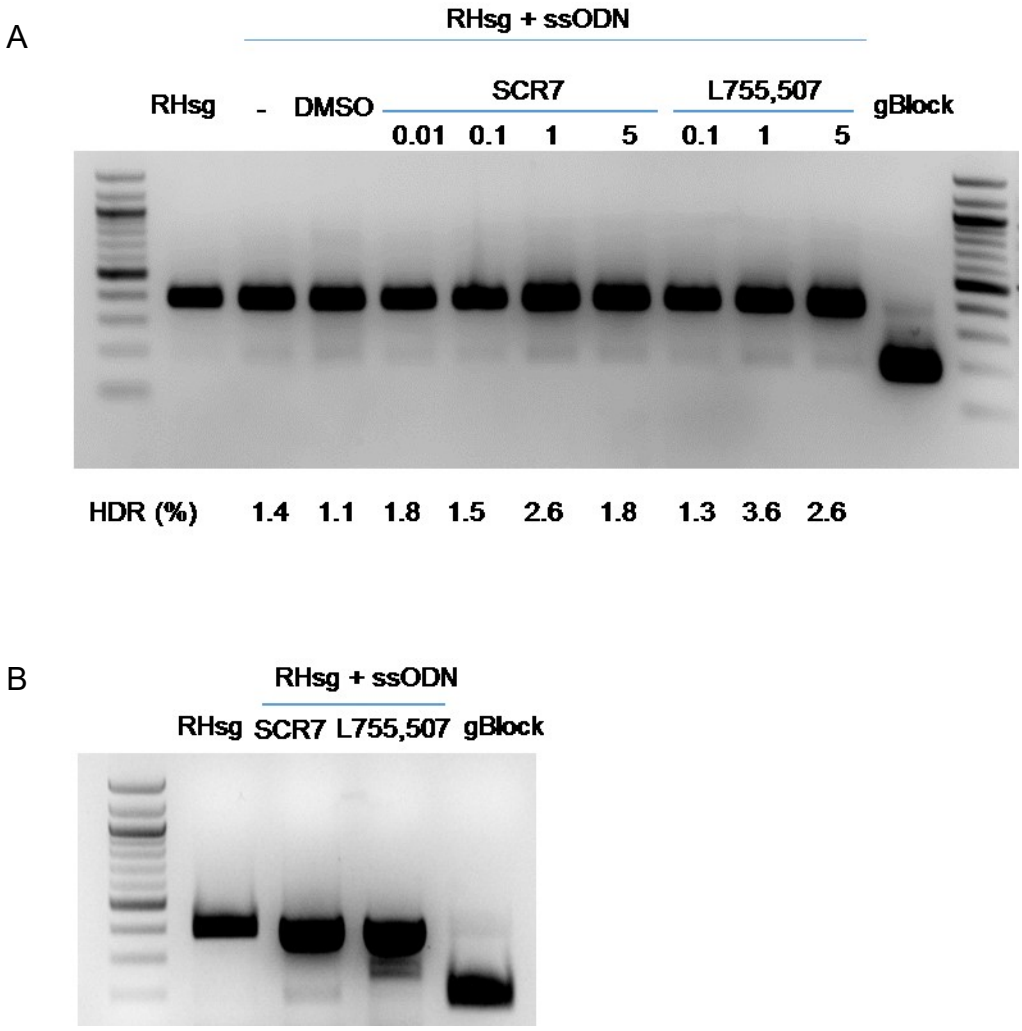


Figure 23. The NHEJ inhibitor Scr7 increases the HDR efficiency.

(A). Different wells of cells were treated with RH sgRNA only (lane 1), RH sgRNA + ssODN (lane 2), RH sgRNA + ssODN + DMSO (lane 3), RH sgRNA + ssODN + Scr7 solved in DMSO of indicated concentration (lane 4-7), and RH sgRNA + ssODN + L755,507 solved in DMSO of indicated concentration (lane 8-10). Amplicons of DNA fragment flanking the targeting site of the genome of each sample and gBlock (positive control) were subjected to restriction enzyme digestion, and HDR efficiency was analyzed. (B) Different wells of cells were treated with RH sgRNA only, RH sgRNA + ssODN +

drug with the concentration showing highest HDR efficiency at panel A. The samples were subjected to restriction enzyme digestion.

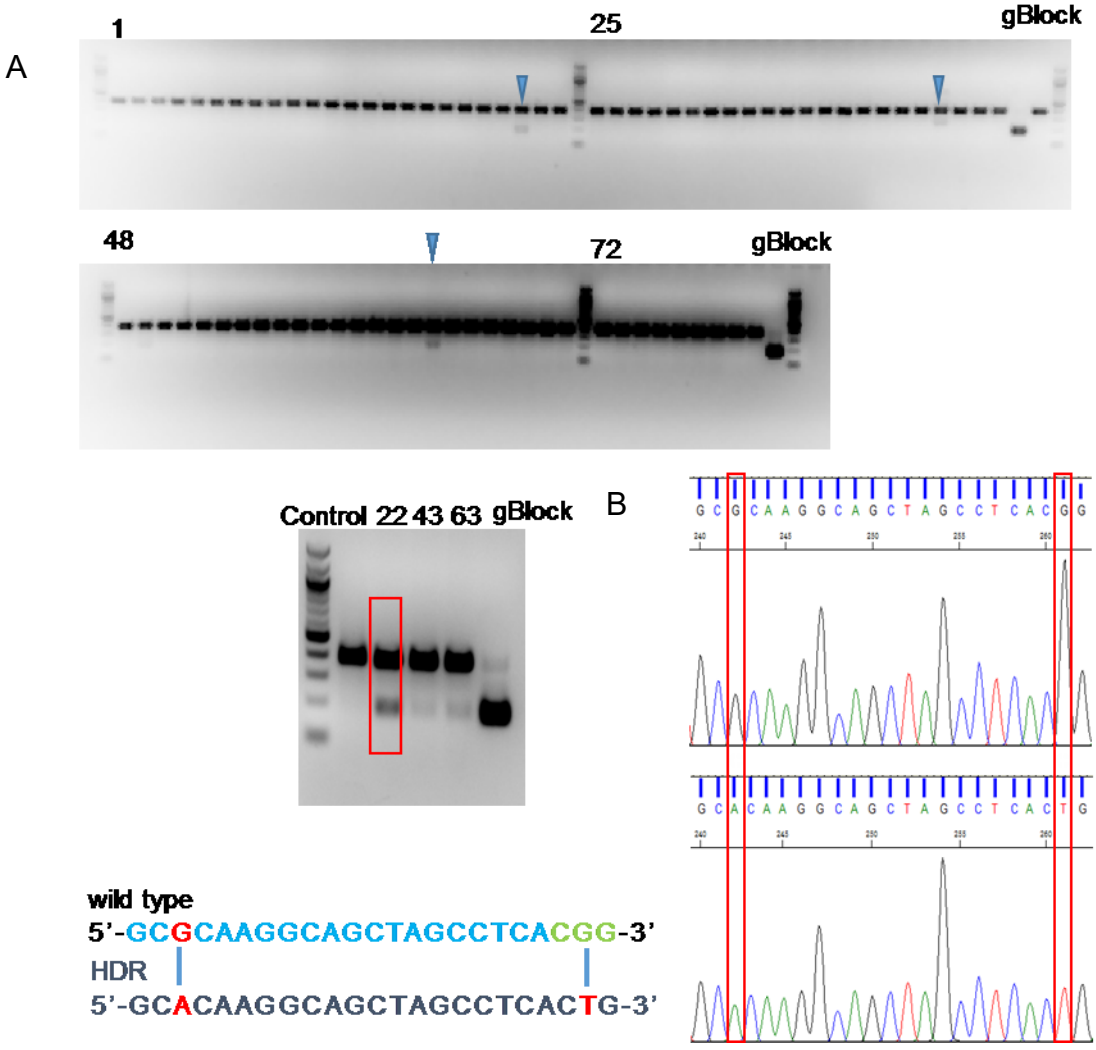


Figure 24. Identify the RH knock-in clones from genotyping by RFLP

(A) Amplicons of DNA fragment flanking the RH sgRNA targeting site of the genome of single clones were subjected to restriction enzyme HpyCH4V digestion for genotyping. (B) Representative chromatogram of Sanger sequencing. Positive clones were confirmed by TA cloning of the amplicon to the plasmid and Sanger sequencing.

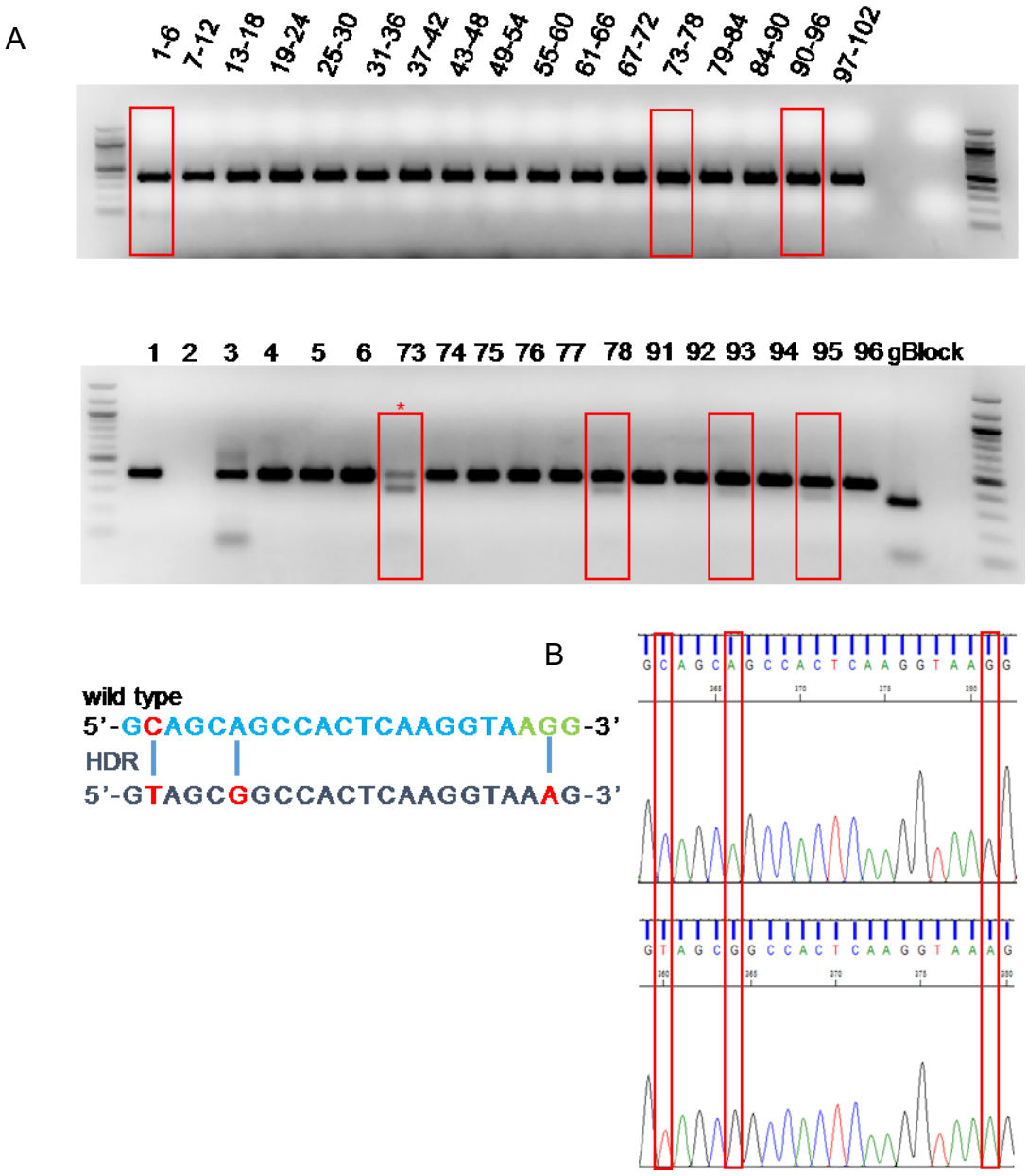


Figure 25. Identify the AV knock-in clones from genotyping by RFLP and confirm by Sanger sequencing.

(A) Amplicons of DNA fragment flanking the AV sgRNA targeting site of the genome of single clones were subjected to restriction enzyme EaeI digestion for genotyping.

(B) Representative chromatogram of Sanger sequencing. Positive clones were confirmed by TA cloning of the amplicon to the plasmid and Sanger sequencing.

Effect of *EIF4G1* PD-associated mutations on protein synthesis

To see if mutations of *EIF4G1* affect global protein synthesis, we chose to use isotope pulse labeling assay to measure the protein synthesis rate. We plated WT, R1205H knock-in, and A502V knock-in cells at sub-confluent status, 24 hour later cells were incubated in cell culture medium containing radioactive ^{35}S - Met/Cys for an hour. Then the cells were lysed, and the protein in part of the lysate was used for liquid scintillation counting, the protein in the rest of the lysate was subject to SDS-PAGE and transferred to a nitrocellulose membrane for phosphorscreen exposure. Radioactive signals in the mutant cells were significantly lower than that in the wild type cells, suggesting a decrease in protein synthesis rate caused by *EIF4G1* mutations. And SDS-PAGE analysis indicated synthesis rates of proteins in a large range of molecular sizes were decreased (Figure 26). Since our transgenic flies showed the PD-like phenotypes, we then wanted to know if the mutations could cause global translation down in the transgenic flies. We crossed the *UAS-WT-EIF4G1, UAS-RH-EIF4G1*, and *UAS-AV-EIF4G1* flies with *ELAV-GAL4* flies, and collected flies with transgene expressed in the whole brain. Then the flies were transferred to ^{35}S - Met/Cys labeled food for 24 hours. The fly heads were harvested and lysed for scintillation counting and SDS-PAGE. Consistent with the observations in the knock-in cells, *EIF4G1* mutant flies showed decreased protein synthesis rate compared with *WT-EIF4G1* transgenic flies (Figure 27). However, we saw similar level of global translation of mutant flies as the Control flies. But we did find the mutant flies showed impaired

phenotypes when compared with the Control flies. We thought the mutant eIF4G1 may affect the subset of the transcripts but the endogenous fly eIF4G compensated the global translation rate.

eIF4G1 is involved in both cap-dependent and cap-independent translation. To test if mutations of *EIF4G1* can affect either or both type of translation initiation, we used a bicistronic reporter assay. As the schematic diagram shows, the bicistronic reporter is composed of a FLAG readout to monitor cap-dependent translation followed by an IRES site with c-Myc readout to monitor cap-independent translation from the same transcript. We transfected the WT, RH, and AV cells with the reporter DNA and harvested cells to analysis the reporter expression. We found both cap-dependent and cap-independent translation were dramatically decreased in AV knock-in cells. The cap-dependent translation in RH cells was slightly decreased, but cap-independent was not. This data reflected the phenotypes we found in the transgenic flies that though both mutations caused PD-like phenotypes, the AV flies always showed the more severe symptoms (Figure 28).

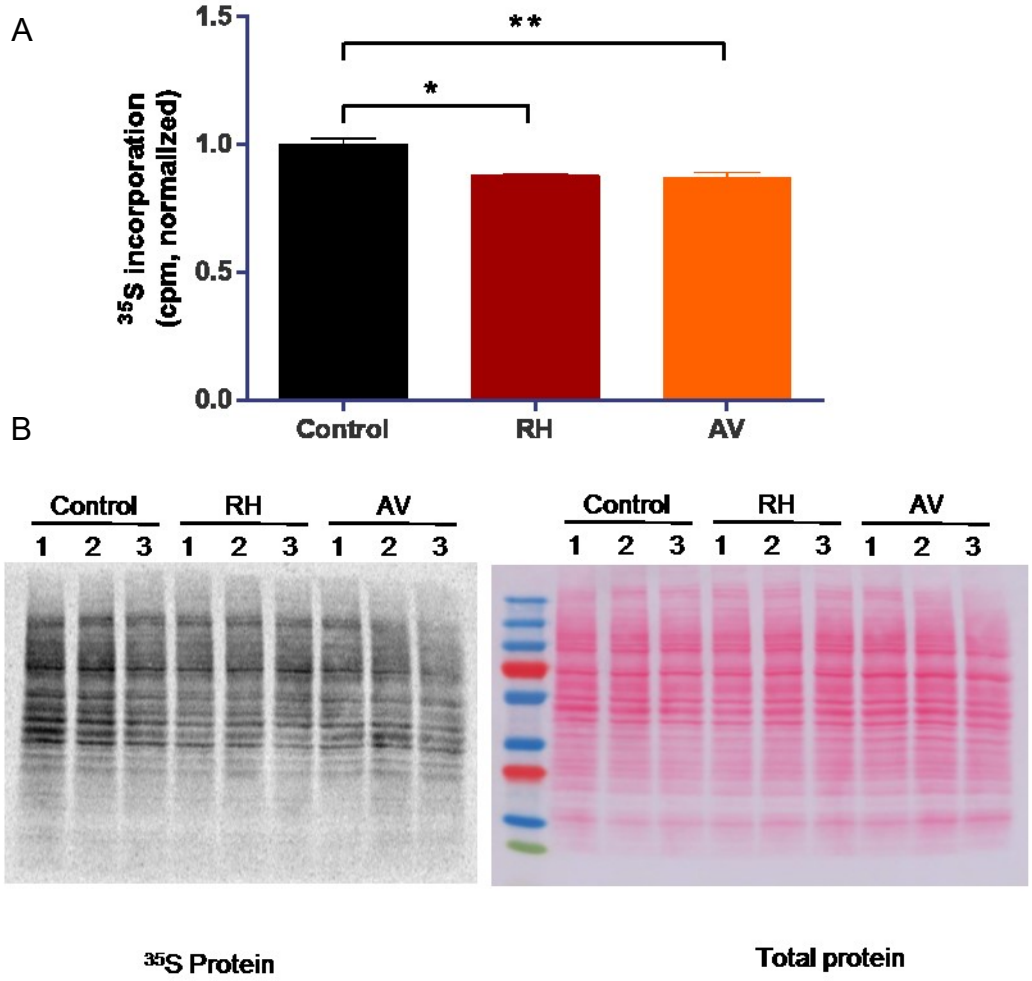


Figure 26. Reduced global translation in RH and AV knock-in cell lines.

Cells in one well of 6-well plate were incubated with 100 μCi of ^{35}S-Met/Cys for one hour. Half of the cell lysate was subjected to SDS-PAGE and the other half was subjected to methanol precipitation for liquid scintillation counting. (A) Protein *de novo* synthesis rate, indicated by the ^{35}S-Met/Cys incorporation measured by liquid scintillation counting, was significantly reduced in RH and AV knock-in cells (ANOVA, Tukey's post hoc test, $^{**}p<0.01$, $^{*}p<0.05$, n=3). (B) Autoradiography from lysates revealed a widespread increase in ^{35}S-Met/Cys-labeled protein abundance. Ponceau staining for total protein.

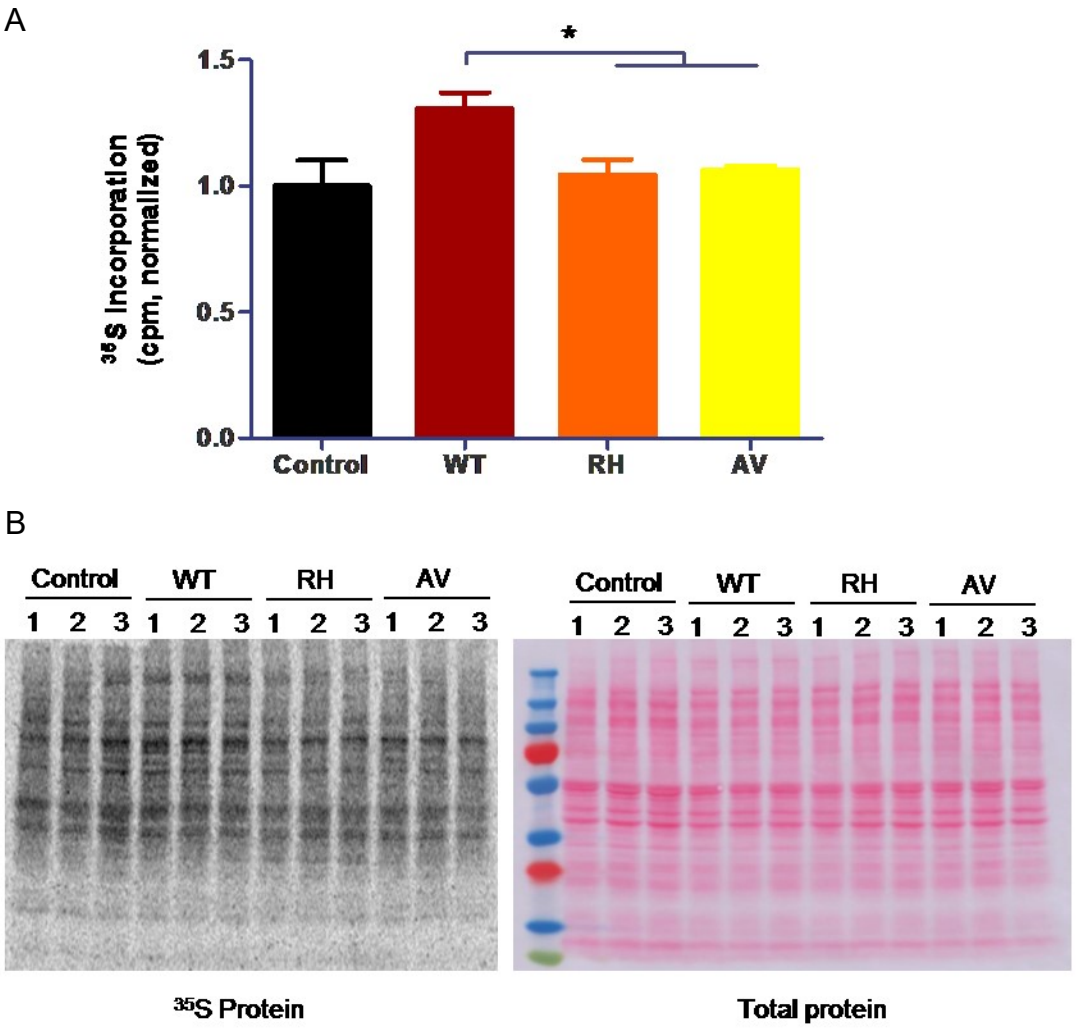


Figure 27. Reduced global translation in RH and AV transgenic flies.

Cohort of 30 flies per vial of 3 vials of each genotype were fed normal food mixed with 100 µCi of ^{35}S-Met/Cys for 24 hours. Fly heads extracts were subjected to SDS-PAGE and scintillation counting. (A) Protein *de novo* synthesis rate was significantly reduced in RH and AV transgenic flies (ANOVA, Tukey's post hoc test, *p<0.05, n=3 vials of 30 flies per vial). (B) Autoradiography from lysates revealed a widespread increase in ^{35}S-Met/Cys-labeled protein abundance. Ponceau staining for total protein.

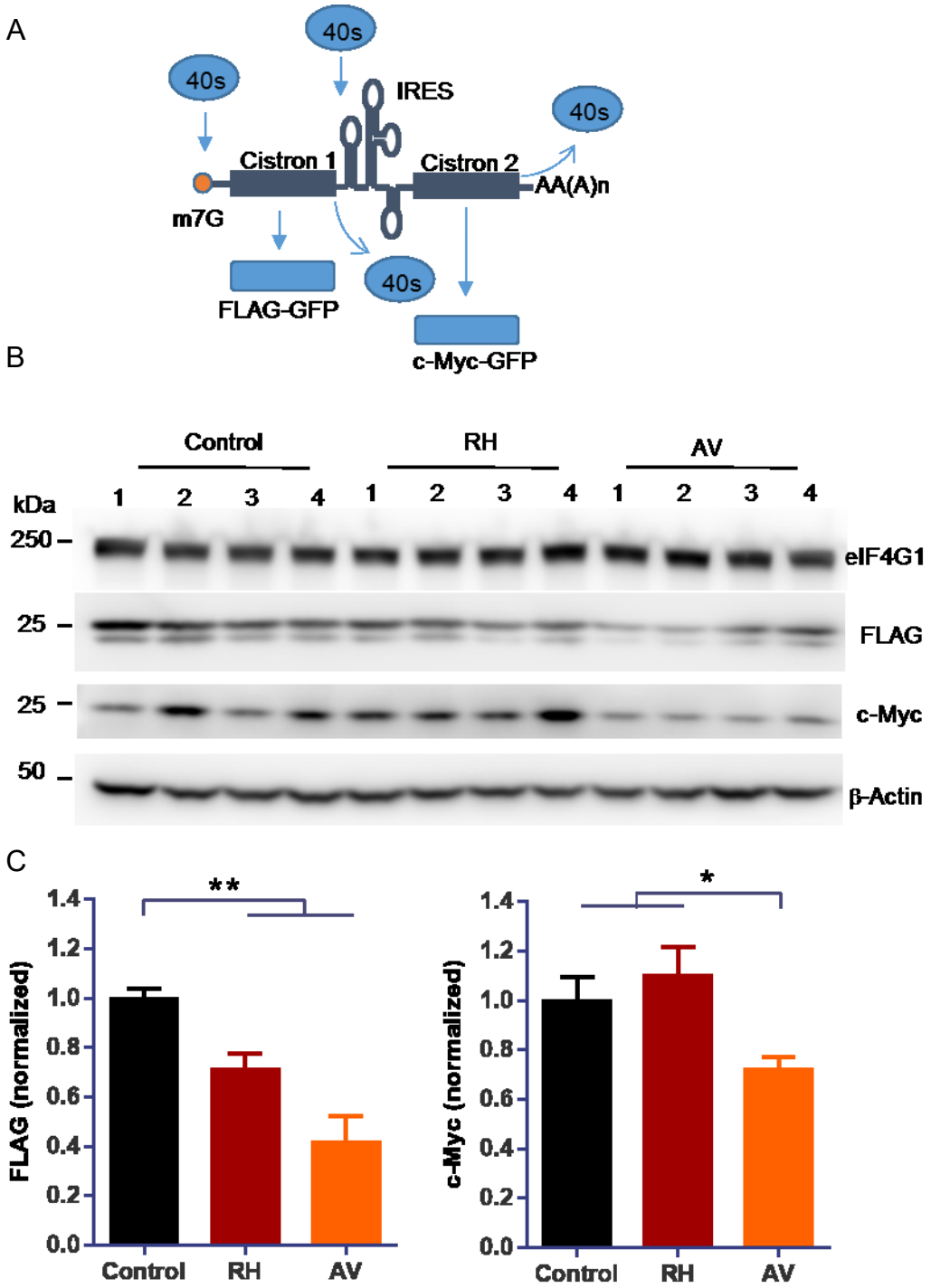


Figure 28. AV mutation reduces both cap-dependent and cap-independent translation, while RH mutation slightly reduces cap-dependent translation.

(A) Schematic representation of the bicistronic reporter used. (B) Representative immunoblots of bi-cistronic reporter assay. Bicistronic reporter was transfected into the WT, RH, AV eIF4G1 knock-in SH-SY5Y cells. After 24hours, cells were collected and lyzed, fractionated by SDS-PAGE, and analyzed by immunoblotting with the indicated antibodies. (C) Quantification of the protein level shows AV mutation reduces both cap-dependent and cap-independent translation, while RH mutation slightly reduces cap-dependent translation (ANOVA, Tukey's post hoc test, $*p<0.05$, n=3)

Discussion

As *EIF4G1* is ubiquitously and abundantly expressed, we realized it was hard to dissect the effects caused by transfected mutant *EIF4G1* in transient transfection system (Figure 14). Thus we have generated *EIF4G1* mutation knock-in cell lines using the genomic editing technology CRISPR/Cas9. When we started to apply this technology, it was less than a year after the first couple of papers just came out. While CRISPR/Cas9 had been reported extensively in genomic engineering to make knock down or knock out, generating knock-in organisms utilizing the Homology-Directed Repair remained inefficient (Yang, Wang *et al.* 2013, Platt, Chen *et al.* 2014, Dow, Fisher *et al.* 2015). There are a few explanations for the low efficiency of HDR editing. HDR occurs only during S and G2 phase, less frequently than NHEJ which could occur at any phase (Lieber 2010). And the nuclease activity of CRISPR/Cas9 is considered very robust, so continuous editing could happen until the targeting site is largely altered (Wang, Yang *et al.* 2013, Canver, Bauer *et al.* 2014). When we were designing the first set of sgRNA and ssODN, we took into consideration that if the targeting site overlaps with the mutation site, the nucleotide change could potentially block the second cleavage (Figure 15). However, we found no positive R1205H knock-in clone by RFLP genotyping out of about 300 clones (Figure 20). By Sanger sequencing of around 300 clones of AVsgRNA edited cells, we identified the intended mutation but companied by undesired indels at several nucleotides distance (Figure 21). We concluded that in our case, single nucleotide in the targeting site was not sufficient to block the further editing.

So we came up with the solution that the continuous cutting could be prevented only if the required PAM sequence is altered. This solution was also mentioned in several later published papers (Hsu, Lander *et al.* 2014, Paquet, Kwart *et al.* 2016). Besides designing a silent blocking mutation of the NGG sequence in our ssODNs, we also introduce a synonymous nucleotide change that created a novel restriction enzyme site for efficient genotyping purpose (Figure 22). To increase the overall efficiency of getting the knock-in clones, we treated the cells with small molecule Scr7, a NHEJ inhibitor (Figure 23). We also applied fluorescence cell sorting to select the cells with Cas9 expressed. Instead of performing single cell isolation right after the sorting, we culture the sorted cells as a whole until they reached stable growth to increase the survival chance of edited cells. And before doing the labor intensive single cell isolation, we genotyped the whole cell population to make sure there were knock-in cells existing and figure out their percentage. We optimized and streamlined the procedure so we were able to generate a knock-in cell line within one month. Both RH and AV knock-in cell lines we got were heterozygous. Since EIF4G1 mutation was identified as autosomal dominant, we did not proceed to make a homozygous line.

As the involvement of eIF4G1in mRNA translation initiation has been well characterized, the first question we wanted to know about the mutations was if they could affect the protein synthesis. ^{35}S-Met/Cys metabolic labeling in the knock-in cells showed the global translation was decreased in mutant cells compared with the wild type cells (Figure 26). It was consistent in transgenic flies that mutations were associated with decreased protein synthesis rate when compared with WT-eIF4G1 expressing flies (Figure

27). These data suggested that *EIF4G1* mutations caused mRNA translation repression, and reduced protein synthesis was associated with neurodegeneration. This finding was a surprise to us, but was not totally unexpected. Our lab reported that increased phosphorylation of ribosomal protein s15 by pathogenic *G2019S LRRK2* elevates global protein synthesis and causes neurotoxicity (Martin, Kim *et al.* 2014). But it was also reported that sustained protein translation repression was associated with synaptic failure and neurodegeneration in prion-diseased mice (Moreno, Radford *et al.* 2012). Also, some other reports showed that decreased protein translation found in age-related neurodegeneration (Douglas and Dillin 2010). It can be explained by that the balance of protein level is crucial for cell survival, either too much or too little is detrimental. Another reason could be different genes regulate different subsets of mRNA transcripts. And survival-required expression levels of different mRNAs are different. In the following studies of phosphorylation of rps15 by LRRK2, our lab did a genome-wide translation profiling with *G2019S LRRK2*, and we found some particular subsets of mRNAs were sensitive to LRRK2 mutation. There were studies about the preference of eIF4G1 on translation initiation of mRNA with particular sequence features. It was reported that in breast cancer cell line, depletion of eIF4G1 would affect the mRNA with multiple upstream open reading frames (Ramirez-Valle, Braunstein *et al.* 2008). There was also a study about that in *C.elegans*, eIF4G1 mediated somatic maintenance by regulating mRNA with longer-than-average ORF length (Rogers, Chen *et al.* 2011). In a recently published paper, eIF4G1 was suggested being required for efficient translation of mRNAs with structured 5' UTR (Liew, Assmann *et al.* 2014).

And we also dissected the effect of the mutations on cap-dependent and cap-independent translation. We found cap-dependent translation was slightly decreased in RH knock-in cells, but cap-independent translation did not. Whereas in AV knock-in cells, both cap-dependent and cap-independent translation were prominently decreased (Figure 28). This finding was consistent with the PD-like phenotypes in the trangenic flies that AV transgenic flies showed more severe symptoms. But we coud not distinguish if the AV mutation abolished the protein function to a larger extend or it affected a wider range of transcripts than the RH mutation. Additional biochemical studies are required to study the effects of the mutations on protein translation in more granular level. The knock-in cell lines will provide a stable platform supporting these studies. And we can also apply the same procedure to generate knock-in embryonic stem cells, and differentiate them to DAergic neurons to study the influence of the mutations on specific neuronal activities.

Materials and Methods

Generation of EIF4G1 RH/AV knock-in SH-SY5Y cell lines

Design sgRNA and ssODNs

EIF4G1 genomic sequence, Chromosome 3: 184,314,495-184,335,358 forward strand. GRCh38:CM000665.2, was acquired from Ensembl. 250 base upstream and 250 bases downstream of the mutation site was used as the input in the online CRISPR Design Tool (http://tools.genome-engineering.org) for generating a list of target sites ranked by off-target scores. The scores and the distance-to-mutation were considered for picking the sgRNA sequences. Due to the requirement of human U6 promoter to have a 'G' base at the transcription start site. A 20bp genome target starting with the base 'G' is preferred, but if the more important criteria can be met by the targeting sequence starting with other bases, an additional 'G' could be added to the front of guide oligo sequence. sgRNA target sequence was put into the following format and synthesized as DNA oligos at Integrated DNA Technologies (IDT).

5' – CACCGNNNNNNNNNNNNNNNNN – 3'
3' – CNNNNNNNNNNNNNNNNNNCAAA – 5'

EIF4G1 genomic sequence 100 bases upstream and downstream of the mutation site was selected for ssODN design. The point mutation nucleotide was changed. The synonymous mutation was made at one of the G sites of NGG PAM. And silent mutation for creating a novel restriction enzyme targeting site around mutation site was identified by the online tool: http://resitefinder.appspot.com/ Both the sense and antisense ssODNs were ordered as 4 nM Ultramer® at IDT.

Generation of sgRNA expression constructs

We used the plasmid-based cloning methods as previously described.

1. The sense and antisense guide oligos were phosphorylated and annealed as following: 1 µL sense oligo (100µM), 1 µL antisense oligo (100µM), 1 µL 10X T4 Ligation Buffer (NEB) 6.5 µL ddH2O, 0.5 µL T4 PNK (NEB) were mixed and annealed in a thermocycler using the following parameters: 37C 30 min, 95C 5 min and then ramp down to 25C at 5C/min.

2. Single-step digestion-ligation.

X µL PX458 (SpCas9(BB)-2A-GFP) vector (100ng), 2 µL phosphorylated and annealed oligo duplex from step 1 (1:250 dilution), 2 µL 10X Tango buffer (Thermo Scientific), 1 µL DTT (10mM to a final concentration of 1mM, Fermentas), 1 µL ATP (10mM to a final concentration of 1mM, New England BioLabs), 1 µL FastDigest *Bbs*I (Thermo Fisher Fermentas), 0.5 µL T7 DNA ligase (New England BioLabs), Y µL ddH2O, 20 µL total were mixed, the ligation reaction was incubated in a thermocycler: 37 C 5 min, 23 C 5 min for 6 cycles, 4 C hold until next step.

3. Treat with PlasmidSafe exonuclease to prevent unwanted recombination products:

11 µL ligation reaction from **step 2,** 1.5 µL 10X PlasmidSafe Buffer (Epicentre), 1.5 µL 10mM ATP (Epicentre), 1 µL PlasmidSafe exonuclease (Epicentre), 15 µL total were mixed, and the reaction was incubated at 37C for 30 min.

4. Transformation with 1-2 μL of the final product into STBL3 competent cells. Pick the colonies and sequence verify the clones.

Cell Culture and Transfection

The human neuroblastoma SH-SY5Y cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum and 100U/ml of penicillin/streptomycin antibiotics and were grown at 37°C, 7% CO2, and 90% humidity. Cells were plated at sub confluent in 6 well plate 24 hours before transfection. Suspended the ssODNs in ddH2O to a final concentration of 10 uM. 500 ng of PX458(sgRNA) was premixed with 1 μL of the ssODN template. Transfections was performed using X-tremeGENE™ transfection reagent according to the manufacturer's instructions (Roche). At the same time, Scr7 (Excess Bioscience) was added to the cell culture medium at the end concentration of 1uM. 48 hours after transfection, cells were harvested for fluorescence-activated cell sorting (FACS). GFP-positive cells were selected, half were cultured, and the rest were for the functional test of genomic editing.

Gene editing efficiency test by the SURVEYOR nuclease assay, the restriction-fragment length polymorphism (RFLP) analysis, and Sanger sequencing

Cells were harvested and lysed in the QuickExtract solution (Epicentre) for genomic DNA extraction according to the manufacturer’s instructions. The final concentration of extracted DNA solution was normalized to 100 ng/ μL with ddH2O. PCR reaction was performed using the high-fidelity polymerase, Herculase II fusion polymerase

(Agilent). Primers were designed to amplify 500 base pair DNA fragment around the sgRNA targeting site.

PCR reaction mix was optimized as following:

Component	Amount(μL)	Final concentration
Herculase II PCR buffer, 5×	10	1x
dNTP, 100 mM (25 mM each)	1	2mM
Fwd primer, 10uM	1.25	0.25uM
Rev primer, 10uM	1.25	0.25uM
DMSO	1	2%
Herculase II fusion polymerase	1	
DNA template	2	4ng/ μL
ddH2O	32.5	
Total	50	

The cycling conditions were also optimized as:

Cycle number	Denature	Anneal	Extend
1	95 C, 2 min		
35		56 C, 30 s	72 C, 30 s
1			72 C, 3 min

5 μL PCR products were used to confirm the experiment worked and the rest was purified with the QIAQuick PCR purification kit (Qiagen).

DNA heteroduplex formation was set up as 2 μΛ of Taq PCR buffer (NEB), 10x, 18 μΛ of Normalized PCR product, 20 ng/μΛ. The annealing reaction was set up the same as the protocol. Then the product was digested by Surveyor® Mutation Detection Kits (IDT) following the manufacturer's instructions. The digestion products were ran at the 2% (wt.vol) agarose gel added with ethidium bromide. The gel was imaged and the intensities of the bands were quantified by ImageJ.

For restriction-fragment length polymorphism (RFLP) analysis, starting with genomic DNA extraction, the following steps were the same before enzyme digestion. The digestion reactions were set up according to the corresponding specific enzyme.

For Sanger sequencing, repeat same steps from DNA extraction to PCR product purification. Then the PCR products were cloned into the plasmid for sequencing using the Zero Blunt® TOPO® PCR Cloning Kit for Sequencing (ThermoFisher Scientific) following the manufacturer's protocols.

Single cell isolation

Passaged the cells for at least two times after cell sorting until the cells were growing stable. Dissociated the cells and pipetted the cell to ensure single cells. Count the cell numbers and performed 10-fold serial dilution to a final concentration of 1 cell per 200 μΛ. Examined the cell concentration at an intermediate stage, adjusted the dilution volume if necessary. Prepared three 96-well plates for each transfection combination. Use mμΛtichannel pipette to add 200 μΛ of cell solution to each well. Monitor the colonies for single cell clonal appearance. Visible clones can be found one week after plating. Allow

the cells to grow for additional one week or two and monitor the confluence. When the cells are more than 80% confluent, dissociate the cells, and use 80% of total cells for genotyping and plate the remaining cells into the replica wells.

De Novo Protein Synthesis measurement by 35S-methionine/cysteine incorporation

SHSY5Y cells were plated at 40%-50% confluent in 6-well plate one night in advance. One milliliter of culture medium was pipetted out into a tube from each well. 35S-methionine/cysteine (PerkinElmer) was added to the medium collection as 100 uCi per well. Then the culture medium mixed with isotope was added back to each well. The plate was placed back into the incubator for one hour. Then took the plate out, removed the radioactive medium, washed the cells with cold PBS for two times. Cells were harvested and lysed in extraction buffer (1% NP-40, 50 mM Tris-HCl, 150 mM NaCl, 5 mM EGTA, protease inhibitor cocktail). Half of the lysate was processed through SDS-PAGE and transferred to nitrocellulose membrane. Total protein amount was visualized by Ponceau S dye (Sigma-Aldrich) staining. Radioactive signal was captured by incubating the membrane in the phosphor screen cassette for 24 hours. The other half of the lysate was concentrated by methanol/heparin protein precipitation. For every 100 μΛ of protein faction, add 1μΛ of Heparin (100 mg/ml), and 400 μΛ methanol, mixed and span at top speed for 5 minutes. Removed the supernatant and dried the pellets in the air. Resuspended the pellets with 20 μl 8M Urea/150 mM Tris (pH 8.5). Added 10 μL of sample into the liquid scintillator, and analyzed the mixture in the Liquid Scintillation Counter. Used 2 μL of the sample to test protein concentration using BCA assay. The final result was the radioactive activity (reported as counts per minute) normalized by the total protein amount.

For protein synthesis measurement in flies, 35S-methionine/cysteine was incorporated into flies through feeding them with food-isotope mixture. Standard solid food was melt and distributed at 1ml per well, 24 hours later, the food was dry. 35S-methionine/cysteine was diluted as 100 uCi per 100 µL ddH2O, then 100 µL of diluted isotope was added to the food. Waited for another 24 hours till the isotope was completely soaked into the food. 30 flies were transferred to the isotope-mixed food and fed for 24 hours. Then the flies were moved to normal food for an hour. Flies were collected into the 1.5ml tube and froze on the dry ice. Fly heads were separated by vortex for 30 seconds and picked manually into a new tube on the ice. Extraction buffer (1% NP-40, 50 mM Tris-HCl, 150 mM NaCl, 5 mM EGTA, protease inhibitor cocktail) was added in the tube and fly heads were homogenized by a motor-driven pestle. Then the tissue was lysed on ice for 30 minutes and centrifuged at 14,000 *g* for 30 min. The supernatant was collected and proceeded as the same method described above.

Bicistronic reporter assay

The bicistronic reporter construct (pCMV-BICEP 4, Sigma), expressing a single transcript induced by the CMV promoter, contains two eGFP open reading frames, one is N-terminal FLAG-tagged, the other is N-terminal c-Myc-tagged. Translation of the upstream FLAG-tagged GFP is cap-dependent, while the ribosome entry site (IRES) derived from the encephalomyo-carditis virus (EMCV) controls the translation of the downstream c-Myc-tagged GFP in a cap-independent manner. SH-SY5Y cells were plated in the 6-well plate at 40% confluent in the previous night. The following day, cells were transfected with 0.5 ug of the bicistronic reporter DNA per well using X-tremeGENE HP

transfection reagent (Roche).24 hr after transfection, cells were harvested and lysed in extraction buffer (1% NP-40, 50 mM Tris-HCl, 150 mM NaCl, 5 mM EGTA, protease inhibitor cocktail). The lysates were resolved on SDS-PAGE gel and subjected to immunoblot analysis.

Materials

Primer oligos and ssODNs

Synthetic oligonucleotides were synthesized by the Integrated DNA Technologies. Sanger sequencing was performed at Genetic Resources Core Facility of Johns Hopkins University School of Medicine

Purpose	Sequence	Template
sgRNA for R1205H	S: CACCGCGCAAGGCAGCTAGCCTCA A: AAACTGAGGCTAGCTGCCTTGCGC	ENSG00000114867
sgRNA for A502V-1	S: CACCGAGCAGCAGCCACTCAAGGTA A: AAACTACCTTGAGTGGCTGCTGCTC	ENSG00000114867
sgRNA for A502V-2	S: CACCGCAACCACACCTTACCTTGAG A: AAACCTCAAGGTAAGGTGTGGTTGC	ENSG00000114867
R1205H knock-in genotyping	F: GTAGCTTGAGCCGAGAACGA R: CTTCTCGCTTCACTGTGGGA	ENSG00000114867
A502V knock-in genotyping	F: CTGAGGAACTGCTCAACGGA R: AATGGGCGGTCATCAAGGAG	ENSG00000114867
Verify the sgRNA cloning in pX458	GAGGGCCTATTTCCCATGATTCC	Pol III promoter U6

Purpose	Sequence
RH ssODN sense	GATCGTGCGCGGACACCTGCTACCAAGCGGAGCTT CAGCAAGGAAGTGGAGGAGCGGAGTAGAGAACGG CCCTCCCAGCCTGAGGGGCTGCACAAGGCAGCTAG CCTCACTGAGGATCGGGACCGTGGGCGGGATGCCG GTGAGAGTCTGGGAGAGGAATGGAGGGAAAGGCA TTGG
RH ssODN anti-sense	CCAATGCCTTTCCCTCCATTCCTCTCCCAGACTCTC ACCGGCATCCCGCCCACGGTCCCGATCCTCAGTGA GGCTAGCTGCCTTGTGCAGCCCCTCAGGCTGGGAG GGCCGTTCTCTACTCCGCTCCTCCACTTCCTTGCTG AAGCTCCGCTTGGTAGCAGGTGTCCGCGCACGATC
AV ssODN sense	GGAGAAGCTGAGAGTGAGAAAGGAGGAGAGGAAC TGCTCCCCCCAGAGAGTACCCCTATTCCAGCCAACT TGTCTCAGAATTTGGAGGCAGTAGCGGCCACTCAA GGTAAAGTGTGGTTGGACGGTAGAGGTAGGGCGGG CTAGGGGATATCGTTCCTTGCCCTCCTTGATGAC
AV ssODN anti-sense	GTCATCAAGGAGGGCAAGGAACGATATCCCCTAGC CCGCCCTACCTCTACCGTCCAACCACACTTTACCTT GAGTGGCCGCTACTGCCTCCAAATTCTGAGACAAG TTGGCTGGAATAGGGGTACTCTCTGGGGGGAGCAG TTCCTCTCCTCCTTTCTCACTCTCAGCTTCTCC

Table 5. Primers and ssODNs used in Chapter 3

All the ssODNs are designed based on the EIF4G1 genomic sequence ENSG00000114867. The ssODNs listed above are the second batch design described in the result session, these were used in the successful generation of the knock-in cell line.

Reagents

All stock solutions for experiments in chapter 3 were prepared following the methods described in the Materials and Methods section in Chapter 2. Other reagents:

30% acrylamide/bis solution, 29:1 (3.3% C) (bio-rad). Restore Western blot stripping buffer (Thermo Scientific). Ponceau S solution (Sigma-Aldrich). Supersignal West Pico/ Femto chemiluminescent substrate (Thermo Scientific). BCA protein assays reagent (Pierce). DNA mini-/ maxi- prep kit (Qiagen). For information regarding components of protease inhibitors and sequence information of primers used in this chapter, refer to Materials and Methods section in Chapter 2.

Antibodies

Conjugant	Antigen	Species	Supplier	Titer	Application
N/A	eIF4G1	rabbit	Cell Signaling	1:1000	WB
HRP	β-Actin	mouse	Sigma-Aldrich	1:10,000	WB
HRP	V5	mouse	Invitrogen	1:5000	WB
HRP	FLAG	mouse	Sigma-Aldrich	1:5,000	WB
HRP	c-Myc	mouse	Invitrogen	1:5000	WB
HRP	mouse IgG	sheep	GE Healthcare	1:2,500	WB
HRP	rabbit IgG	donkey	GE Healthcare	1:2,500	WB
Alexa Fluor 488	mouse IgG(H+L)	donkey	Invitrogen	1:500	IF

Table 6. Information of antibodies used in Chapter 3

Statistical Analysis

Statistical significance was determined by one-way Analysis of variance (ANOVA) and Tukey's post hoc test for comparison among multiple groups of more than three, or a two-tailed nonpaired Student t-test for comparison of two groups (control and test) unless otherwise noted in the figure legends. The highest significance level was set at $p<0.05$

Chapter 4

Conclusions

Missense mutation c.3614G>A (p.R1205H) of *EIF4G1* gene has been identified in autosomal dominant familial Parkinson disease cases (Chartier-Harlin, Dachsel *et al.* 2011). The c.1505C>T (p.Ala502Val) mutation was the most frequent substitution identified in some small families in the same study, and also suggested from an ancestral founder in the haplotype analysis. Following studies have also identified other variants (Figure 1), but the mutations of *EIF4G1* gene seemed to be rare, and genetic studies were not enough to prove the mutations were causal (Deng, Wu *et al.* 2015). However, we believe the study of *EIF4G1* mutations will help us understand better about PD pathogenesis, as eIF4G1 is crucial in protein translation and many studies have revealed that aberrant protein translation can lead to various neurological disorders, including neurodegenerative disease. To answer if and how the mutations of *EIF4G1* could lead to Parkinson disease, we generated a transgenic *Drosophila* model and knock-in SH-SY5Y cell lines. We found that overexpression of RH-eIF4G1 and AV-eIF4G1 in dopaminergic neurons of the flies caused PD-like phenotypes, and mutant flies showed the less global translation than the WT-eIF4G1 expressing flies, but the same as the control flies. With the abundant amount of endogenous eIF4G in the transgenic model, we cannot tell whether there was compensation effect or the mutations were dominant negative. Later we found both mutations suppressed the global protein synthesis in knock-in cells.

Development of transgenic *Drosophila* for *EIF4G1*-linked Parkinson's disease modeling

By employing the *GAL4*/*UAS* system, we have generated tissue-specific *WT-EIF4G1*, *R1205H-EIF4G1*, and *A502V-EIF4G1* transgenic flies. As the *TH-GAL4* line, in which *GAL4* is regulated under the promoter of *Drosophila tyrosine hydroxylase (DTH)* gene (Friggi-Grelin, Coulom *et al.* 2003), was already available in the Bloomington *Drosophila* Stock Center and proved to induce robust *GAL4* expression specifically in dopaminergic neurons, our main effort was focused on optimizing the *UAS* line.

As we wanted to compare overexpression of wild type versus mutants, the equal levels of protein expression across the transgenic lines were critical. And eIF4G1 is abundantly and ubiquitously expressed, so the transgene expression level must be high enough to overcome the effect of endogenous wild type protein. We have tried three different types of constructs for generating the lines. We started with the traditional *P*-element induced transformation. As the "position effects" of random insertion of *P*-element into the chromatin, both the mRNA and protein levels of the transgene across the lines were largely varied (Figure 4, 5). It is accepted that if the *P*-element induced transformation lines are used, there should be at least two independent lines from each genotype expressing the similar level of the transgene. So to overcome the problem of the transgene expression variance, we switched to the site-specific integration system induced by the phiC31 integrase (Groth, Fish *et al.* 2004), a bacteriophage isolated integrase that could mediate the unidirectional sequence-specific recombination between two DNA fragments, *attB* and *attP* (Figure 6). Konrad Basler's group developed the host lines that have both *attP* site

and PhiC31 integrase engineered into the genome, so *attB* DNA constructs can be inserted at the known genomic location (Bischof, Maeda *et al.* 2007). However, we found although the transgene expression levels across the PhiC31 lines were equal, the absolute level was low compared with the *P*-element lines, so we had to find the new strategy that allowed us to have a broader range of transgene expression level. Rubin Gerald's group modified the *attB* plasmid by adding the copy numbers of *UAS* from 5 to 10, 20, and 40 and developed the *pJFRC* plasmid series (Figure 7) (Pfeiffer, Ngo *et al.* 2010). We cloned the *EIF4G1* coding sequence into the *pJFRC-10xUAS* and *pJFRC-20xUAS* and received the lines expressing robust and equal levels of the transgene (Figure 9).

For characterization of the transgenic models, we took the mortality record, tested the motor ability, and examined the neurodegeneration. Life expectancy of patients with Parkinson disease is reduced compared with the general population. We found expression of either *RH-EIF4G1* or *AV-EIF4G1* caused premature mortality compared with WT-EIF4G1 expressing flies. The mean lifespans of *RH-EIF4G1* and *AV-EIF4G1* transgenic flies survived were 54 and 40 days, respectively, while the mean lifespan of *WT-EIF4G1* flies is 62 days. The ages at which 50% of the *RH-EIF4G1* and *AV-EIF4G1* transgenic flies died were 54 and 39 days, respectively, compared with 65 days of *WT-EIF4G1* and non-transgenic control flies (Figure 10). Interestingly, though the R1205H mutation was the one that identified in the P30 family, AV mutation caused more reduced life span.

As PD is typically diagnosed and characterized by motor dysfunctions, we tested the locomotion ability of the transgenic flies by climbing assays. We found that at 5-week age, the performances of the *RH-EIF4G1* and *AV-EIF4G1* transgenic flies were

significantly worse than WT transgenic and Control flies, though at early age there was no difference between the flies expressing mutant *EIF4G1* and *WT-EIF4G1* (Figure 11). Moreover, *AV-EIF4G1* flies showed more severe motor defects than *RH-EIF4G1* flies (Figure 11).

The most convincing proof that the mutations are causal is the neurodegeneration in the flies, as the pathological hallmark of PD is the loss of dopaminergic neurons in the substantia nigra. We stained the TH neurons in both young and aged flies, and found at 6-week age, the TH-positive neurons in PPL1, PPM1/2 clusters of *RH-EIF4G1* and *AV-EIF4G1* flies were significantly less than those of *WT-EIF4G1* and Control flies (Figure 13). And there was DAergic neurons loss in PPM3 of *AV-EIF4G1* flies compared with *WT-EIF4G1* and Control flies, while there was no significant difference among all the lines when they were young (Figure 12).

Our data suggested overexpression of mutant *EIF4G1* in dopaminergic neurons could result in neuronal loss, motor function impairment, and early mortality. Several questions remained unanswered. First, we have not tested if the neurodegeneration is dopaminergic neuron specific. It could be addressed by crossing the *UAS* lines with either *dopa decarboxylase (ddc)-GAL4* driver line (Feany and Bender 2000), in which *GAL4* is expressed in both TH and serotonin (5-HT) neurons, or *embryonic lethal abnormal visual system (elav)-GAL4* line (Koushika, Lisbin *et al.* 1996), in which *GAL4* expression is pan-neuronal. But different *GAL4* lines have different driver strength, and the expression level of the transgene is critical in our case, it will be tricky to adjust the extent of the transgene expression. Second, the A502V mutation seemed to cause more severe symptoms than the

R1205H mutation. As the two mutations are far from each other in the primary sequence, they may have different effects on the protein function. We found one of the functional differences of the two mutations in the mechanism studies, and this will be discussed later.

Generation of *EIF4G1* mutations knock-in cell lines for mechanistic study

As function of eIF4G1 has been well studied in protein translation initiation, and numerous studies have suggested the role of translation in neurodegeneration, so to study the mechanism of how the mutations of *EIF4G1* result in neurodegeneration, we firstly wanted to know if the mutations affect protein translation initiation. For *in vitro* studies, we started with transient transfection system. But we hardly saw overexpression of eIF4G1 in transfected cells while V5 tag immunoblotting confirmed the expression of transfected eIF4G1 (Figure 14). We assumed it was because the endogenous eIF4G1 was abundantly expressed and the overall eIF4G1 protein amount was highly regulated, it was hard to induce strong exogenous eIF4G1 expression. Without knowing how much portion of the overall eIF4G1 pool was taken by transfected mutant eIF4G1, it was difficult to dissect the role of the mutations. And since cell plating confluency, transfection reagents toxicity, and transfection efficiency all have an impact on protein translation, so results of the in vitro studies were inconsistent and unreliable. So we pivoted to generating *RH-EIF4G1* and AV-EIF4G1 knock-in cell lines by using the CRISPR/Cas9 system.

At the time we adopted CRISPR/Cas9 in the lab, the majority of the studies reported were making knock-down or knock-out organisms, but generating knock-in systems taking advantage of the cellular homology-directed repair mechanism remained inefficient (Wang, Yang *et al.* 2013, Yang, Wang *et al.* 2013, Platt, Chen *et al.* 2014, Dow, Fisher *et al.* 2015).

One potential idea was about the nuclease activity of CRISPR/Cas9 is so robust that continuous editing could happen until the targeting site is largely altered (Wang, Yang *et al.* 2013, Canver, Bauer *et al.* 2014, Paquet, Kwart *et al.* 2016). So for our first design of single-guide RNA (sgRNA) and single-stranded DNA oligonucleotides (ssODNs) repair template, we considered that if the targeting site overlaps with the mutation site, the nucleotide change could potentially block the second cleavage (Figure 15). Though our sgRNAs and ssODNs were functional and efficient in HEK cells (Figure 16-29), we did not get positive knock-in clones in both RH and AV sgRNAs treated cells (Figure 20, 21). By sequencing the clones, we found unwanted indel mutations presented together with HDR events (Figure 21). We realized the re-cut occurred at the edited site even though our chosen sgRNA targeting sequence overlapped with the intended mutation site. This indicated that one nucleotide difference might not be sufficient to block the binding of sgRNA with the genomic DNA, so the second or more cleavages continued to happen until the NHEJ-mediated modification prevented the further cutting. So in our second design, we induced a synonymous mutation at the binding-required NGG protospacer adjacent motif (PAM) sequence. And to ensure large scale genotyping, we also designed a silent mutation at the repair template to create a restriction enzyme targeting site if the intended mutation itself did not present to be a restriction enzyme targeting site (Figure 22). We also tried the cell sorting enrichment and NHEJ inhibitor treatment to increase the chance of identifying positive knock-in clones (Figure 23). We eventually got the knock-in cell lines of both mutations, and sequencing results confirmed both lines were heterozygous (Figure 24, 25). Through this process, we have optimized the procedure of generating knock-in cell

lines using CRISPR-Cas9. We can apply the same protocol to generate knock-in embryonic stem cells, which can be differentiated into dopaminergic neurons.

We used a radioactive ^{35}S-Met/Cys labeling assay to measure the protein synthesis in the wild type and mutation knock-in cells. Liquid scintillation reads of the radioactive signal in the mutant cells were significantly lower than that in the wild-type cells, suggesting a decrease in protein synthesis rate. And SDS-PAGE analysis of the whole cell lysate indicated protein synthesis in a large range of molecular size was decreased (Figure 26). And after feeding the flies with ^{35}S-Met/Cys mixed food, we also detected the decreased protein synthesis rate of mutant *EIF4G1* flies compared with that of *WT-EIF4G1* transgenic flies (Figure 27). These data suggested that *EIF4G1* mutations were either dominant negative or affecting subset of transcripts, and potential altered protein synthesis was associated with neurodegeneration. This finding was surprising to us, particularly because our lab just reported pathogenic *G2019S LRRK2* could elevate global protein synthesis and cause neurotoxicity. But it was also reported that sustained protein translation repression was associated with synaptic failure and neurodegeneration in prion-diseased mice. It could be explained that protein homeostasis is crucial for cell survival, either too much or too little is detrimental.

Also, bicistronic reporter assays dissecting cap-dependent and cap-independent translation showed both cap-dependent and cap-independent translation were dramatically decreased in AV knock-in cells, while slightly reduced cap-dependent translation in RH knock-in cells (Figure 28). These finding helped us understand better about AV transgenic flies showed more severe symptoms compared with RH transgenic flies. It could due to the

dose-dependent effect on cap-dependent translation as well as AV mutation particularly affects the cap-independent translation. Though only 10-15% of the total mRNAs are translated by the cap-independent manner, which could partially explain the little difference about the effect on global translation of RH and AV, these mRNAs are particularly translated in the conditions such as cellular stress states, viral infection, and diseases. So the hypothesis could be as the flies were aging and accumulating stresses, the AV mutation-bearing flies were more vulnerable to the stress conditions, thus they developed worse outcomes.

The next question we want to answer is why the mutations could cause dopaminergic neuron specific loss. Though at this stage, we know the mutations reduce global translation, we have to study at more granular level to address the specificity question. Translation initiation factors seem to function in a general manner, their preference for specific mRNAs features under particular conditions has been discovered and realized. There were studies about the preference of eIF4G1 on translation initiation of mRNA with particular sequence features. It was reported that in breast cancer cell line, depletion of eIF4G1 would affect the mRNA with multiple upstream open reading frames (Ramirez-Valle, Braunstein *et al.* 2008). There was also a study about that in *C.elegans*, eIF4G1 mediated somatic maintenance by regulating mRNA with longer-than-average ORF length (Rogers, Chen *et al.* 2011). In a recently published paper, eIF4G1 was suggested being required for efficient translation of mRNAs with structured 5' UTR (Liew, Assmann *et al.* 2014). We hypothesize that the mutations could alter the preference of eIF4G1 on particular mRNA sequence features like coding sequence lengths, the presence

of upstream open reading frames (uORF) and upstream start codon (uATG) and secondary structures in the 5′ untranslated region (UTR) and so on. So our next steps will be employing the genome-wide ribosome-profiling technique to identify the differential translated mRNAs due to mutations in *EIF4G1* and further analyze the sequence features of theses mRNAs in the transgenic fly model and *EIF4G1* mutation knock-in SH-SY5Y cells and even hES cell-derived DAergic neurons. So we will be able to expand our knowledge of the specific genes and signal pathways that lead to Parkinson disease and provide insights into the novel therapeutic targets.

Bibliography

Abbas, N., C. B. Lucking, S. Ricard, A. Durr, V. Bonifati, G. De Michele, S. Bouley, J. R. Vaughan, T. Gasser, R. Marconi, E. Broussolle, C. Brefel-Courbon, B. S. Harhangi, B. A. Oostra, E. Fabrizio, G. A. Bohme, L. Pradier, N. W. Wood, A. Filla, G. Meco, P. Denefle, Y. Agid and A. Brice (1999). "A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease." Hum Mol Genet **8**(4): 567-574.

Ali, I. K., L. McKendrick, S. J. Morley and R. J. Jackson (2001). "Truncated initiation factor eIF4G lacking an eIF4E binding site can support capped mRNA translation." Embo j **20**(15): 4233-4242.

Ali, Y. O., W. Escala, K. Ruan and R. G. Zhai (2011). "Assaying locomotor, learning, and memory deficits in Drosophila models of neurodegeneration." J Vis Exp(49).

Andreou, A. Z. and D. Klostermeier (2013). "The DEAD-box helicase eIF4A: paradigm or the odd one out?" RNA Biol **10**(1): 19-32.

Barone, P. (2010). "Neurotransmission in Parkinson's disease: beyond dopamine." Eur J Neurol **17**(3): 364-376.

Bauer, C., I. Diesinger, N. Brass, H. Steinhart, H. Iro and E. U. Meese (2001). "Translation initiation factor eIF-4G is immunogenic, overexpressed, and amplified in patients with squamous cell lung carcinoma." Cancer **92**(4): 822-829.

Bellen, H. J., R. W. Levis, G. Liao, Y. He, J. W. Carlson, G. Tsang, M. Evans-Holm, P. R. Hiesinger, K. L. Schulze, G. M. Rubin, R. A. Hoskins and A. C. Spradling (2004). "The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes." Genetics **167**(2): 761-781.

Bhaya, D., M. Davison and R. Barrangou (2011). "CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation." Annu Rev Genet **45**: 273-297.

Bischof, J., R. K. Maeda, M. Hediger, F. Karch and K. Basler (2007). "An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases." Proc Natl Acad Sci U S A **104**(9): 3312-3317.

Blanckenberg, J., C. Ntsapi, J. A. Carr and S. Bardien (2014). "EIF4G1 R1205H and VPS35 D620N mutations are rare in Parkinson's disease from South Africa." Neurobiol Aging **35**(2): 445.e441-443.

Bonifati, V., P. Rizzu, F. Squitieri, E. Krieger, N. Vanacore, J. C. van Swieten, A. Brice, C. M. van Duijn, B. Oostra, G. Meco and P. Heutink (2003). "DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism." Neurol Sci **24**(3): 159-160.

Braak, H., K. Del Tredici, U. Rub, R. A. de Vos, E. N. Jansen Steur and E. Braak (2003). "Staging of brain pathology related to sporadic Parkinson's disease." Neurobiol Aging **24**(2): 197-211.

Brand, A. H. and N. Perrimon (1993). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." Development **118**(2): 401-415.

Canver, M. C., D. E. Bauer, A. Dass, Y. Y. Yien, J. Chung, T. Masuda, T. Maeda, B. H. Paw and S. H. Orkin (2014). "Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells." J Biol Chem **289**(31): 21312-21324.

Castro, J. P. and C. M. Carareto (2004). "Drosophila melanogaster P transposable elements: mechanisms of transposition and regulation." Genetica **121**(2): 107-118.

Chartier-Harlin, M. C., J. C. Dachsel, C. Vilarino-Guell, S. J. Lincoln, F. Lepretre, M. M. Hulihan, J. Kachergus, A. J. Milnerwood, L. Tapia, M. S. Song, E. Le Rhun, E. Mutez, L. Larvor, A. Duflot, C. Vanbesien-Mailliot, A. Kreisler, O. A. Ross, K. Nishioka, A. I. Soto-Ortolaza, S. A. Cobb, H. L. Melrose, B. Behrouz, B. H. Keeling, J. A. Bacon, E. Hentati, L. Williams, A. Yanagiya, N. Sonenberg, P. J. Lockhart, A. C. Zubair, R. J. Uitti, J. O. Aasly, A. Krygowska-Wajs, G. Opala, Z. K. Wszolek, R. Frigerio, D. M. Maraganore, D. Gosal, T. Lynch, M. Hutchinson, A. R. Bentivoglio, E. M. Valente, W. C. Nichols, N. Pankratz, T. Foroud, R. A. Gibson, F. Hentati, D. W. Dickson, A. Destee and M. J. Farrer (2011). "Translation initiator EIF4G1 mutations in familial Parkinson disease." Am J Hum Genet **89**(3): 398-406.

Chartier-Harlin, M. C., J. Kachergus, C. Roumier, V. Mouroux, X. Douay, S. Lincoln, C. Levecque, L. Larvor, J. Andrieux, M. Hulihan, N. Waucquier, L. Defebvre, P. Amouyel, M. Farrer and A. Destee (2004). "Alpha-synuclein locus duplication as a cause of familial Parkinson's disease." Lancet **364**(9440): 1167-1169.

Chaudhuri, K. R., D. G. Healy and A. H. Schapira (2006). "Non-motor symptoms of Parkinson's disease: diagnosis and management." Lancet Neurol **5**(3): 235-245.

Chen, F., S. M. Pruett-Miller, Y. Huang, M. Gjoka, K. Duda, J. Taunton, T. N. Collingwood, M. Frodin and G. D. Davis (2011). "High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases." Nat Methods **8**(9): 753-755.

Chen, Y., K. Chen, W. Song, X. Chen, B. Cao, R. Huang, B. Zhao, X. Guo, J. Burgunder, J. Li and H. F. Shang (2013). "VPS35 Asp620Asn and EIF4G1 Arg1205His mutations are rare in Parkinson disease from southwest China." Neurobiol Aging **34**(6): 1709.e1707-1708.

Cho, S. W., S. Kim, J. M. Kim and J. S. Kim (2013). "Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease." Nat Biotechnol **31**(3): 230-232.

Cleary, J. D. and L. P. W. Ranum (2013). Repeat-associated non-ATG (RAN) translation in neurological disease. Human Molecular Genetics, Oxford University Press. **22:** R45-51.

Coldwell, M. J., U. Sack, J. L. Cowan, R. M. Barrett, M. Vlasak, K. Sivakumaran and S. J. Morley (2012). "Multiple isoforms of the translation initiation factor eIF4GII are generated via use of alternative promoters, splice sites and a non-canonical initiation codon." Biochem J **448**(1): 1-11.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." Science **339**(6121): 819-823.

Conway, K. A., S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson and P. T. Lansbury, Jr. (2000). "Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy." Proc Natl Acad Sci U S A **97**(2): 571-576.

Dawson, T. M. and V. L. Dawson (2010). "The role of parkin in familial and sporadic Parkinson's disease." Mov Disord **25 Suppl 1**: S32-39.

Dawson, T. M., H. S. Ko and V. L. Dawson (2010). "Genetic animal models of Parkinson's disease." Neuron **66**(5): 646-661.

Deng, H., W. Le, Y. Guo, C. B. Hunter, W. Xie and J. Jankovic (2005). "Genetic and clinical identification of Parkinson's disease patients with LRRK2 G2019S mutation." Ann Neurol **57**(6): 933-934.

Deng, H., Y. Wu and J. Jankovic (2015). "The EIF4G1 gene and Parkinson's disease." Acta Neurol Scand **132**(2): 73-78.

Deveau, H., J. E. Garneau and S. Moineau (2010). "CRISPR/Cas system and its role in phage-bacteria interactions." Annu Rev Microbiol **64**: 475-493.

Dhungel, N., S. Eleuteri, L. B. Li, N. J. Kramer, J. W. Chartron, B. Spencer, K. Kosberg, J. A. Fields, K. Stafa, A. Adame, H. Lashuel, J. Frydman, K. Shen, E. Masliah and A. D. Gitler (2015). "Parkinson's disease genes VPS35 and EIF4G1 interact genetically and converge on alpha-synuclein." Neuron **85**(1): 76-87.

Douglas, P. M. and A. Dillin (2010). "Protein homeostasis and aging in neurodegeneration." J Cell Biol **190**(5): 719-729.

Dow, L. E., J. Fisher, K. P. O'Rourke, A. Muley, E. R. Kastenhuber, G. Livshits, D. F. Tschaharganeh, N. D. Socci and S. W. Lowe (2015). "Inducible in vivo genome editing with CRISPR-Cas9." Nat Biotechnol **33**(4): 390-394.

Elliott, B., C. Richardson, J. Winderbaum, J. A. Nickoloff and M. Jasin (1998). "Gene conversion tracts from double-strand break repair in mammalian cells." Mol Cell Biol **18**(1): 93-101.

Elliott, D. A. and A. H. Brand (2008). "The *GAL4* system : a versatile system for the expression of genes." Methods Mol Biol **420**: 79-95.

Fahn, S. (2003). "Description of Parkinson's disease as a clinical syndrome." Ann N Y Acad Sci **991**: 1-14.

Feany, M. B. and W. W. Bender (2000). "A Drosophila model of Parkinson's disease." Nature **404**(6776): 394-398.

Foeger, N., E. Kuehnel, R. Cencic and T. Skern (2005). "The binding of foot-and-mouth disease virus leader proteinase to eIF4GI involves conserved ionic interactions." Febs j **272**(10): 2602-2611.

Friggi-Grelin, F., H. Coulom, M. Meller, D. Gomez, J. Hirsh and S. Birman (2003). "Targeted gene expression in Drosophila dopaminergic cells using regulatory sequences from tyrosine hydroxylase." J Neurobiol **54**(4): 618-627.

Fujioka, S., C. Sundal, A. J. Strongosky, M. C. Castanedes, R. Rademakers, O. A. Ross, C. Vilarino-Guell, M. J. Farrer, Z. K. Wszolek and D. W. Dickson (2013). "Sequence variants in eukaryotic translation initiation factor 4-gamma (eIF4G1) are associated with Lewy body dementia." Acta Neuropathol **125**(3): 425-438.

Gagliardi, M., G. Annesi, P. Tarantino, G. Nicoletti and A. Quattrone (2014). "Frequency of the ASP620ASN mutation in VPS35 and Arg1205His mutation in EIF4G1 in familial Parkinson's disease from South Italy." Neurobiol Aging **35**(10): 2422.e2421-2422.

Gallie, D. R. (1991). "The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency." Genes Dev **5**(11): 2108-2116.

Gargano, J. W., I. Martin, P. Bhandari and M. S. Grotewiel (2005). "Rapid iterative negative geotaxis (RING): a new method for assessing age-related locomotor decline in Drosophila." Exp Gerontol **40**(5): 386-395.

Garneau, J. E., M. E. Dupuis, M. Villion, D. A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A. H. Magadan and S. Moineau (2010). "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." Nature **468**(7320): 67-71.

Gasiunas, G., R. Barrangou, P. Horvath and V. Siksnys (2012). "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria." Proc Natl Acad Sci U S A **109**(39): E2579-2586.

Gkogkas, C. G., A. Khoutorsky, I. Ran, E. Rampakakis, T. Nevarko, D. B. Weatherill, C. Vasuta, S. Yee, M. Truitt, P. Dallaire, F. Major, P. Lasko, D. Ruggero, K. Nader, J.-C. Lacaille and N. Sonenberg (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. Nature. **493:** 371-377.

Golbe, L. I., G. Di Iorio, V. Bonavita, D. C. Miller and R. C. Duvoisin (1990). "A large kindred with autosomal dominant Parkinson's disease." Ann Neurol **27**(3): 276-282.

Goldwurm, S., A. Di Fonzo, E. J. Simons, C. F. Rohe, M. Zini, M. Canesi, S. Tesei, A. Zecchinelli, A. Antonini, C. Mariani, N. Meucci, G. Sacilotto, F. Sironi, G. Salani, J. Ferreira, H. F. Chien, E. Fabrizio, N. Vanacore, A. Dalla Libera, F. Stocchi, C. Diroma, P. Lamberti, C. Sampaio, G. Meco, E. Barbosa, A. M. Bertoli-Avella, G. J. Breedveld, B. A. Oostra, G. Pezzoli and V. Bonifati (2005). "The G6055A (G2019S) mutation in LRRK2 is frequent in both early and late onset Parkinson's disease and originates from a common ancestor." J Med Genet **42**(11): e65.

Groth, A. C., M. Fish, R. Nusse and M. P. Calos (2004). "Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31." Genetics **166**(4): 1775-1782.

Haghighat, A. and N. Sonenberg (1997). "eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure." J Biol Chem **272**(35): 21677-21680.

Hasty, P., J. Rivera-Perez and A. Bradley (1991). "The length of homology required for gene targeting in embryonic stem cells." Mol Cell Biol **11**(11): 5586-5591.

Hatano, Y., Y. Li, K. Sato, S. Asakawa, Y. Yamamura, H. Tomiyama, H. Yoshino, M. Asahina, S. Kobayashi, S. Hassin-Baer, C. S. Lu, A. R. Ng, R. L. Rosales, N. Shimizu, T. Toda, Y. Mizuno and N. Hattori (2004). "Novel PINK1 mutations in early-onset parkinsonism." Ann Neurol **56**(3): 424-427.

Heman-Ackah, S. M., M. Hallegger, M. S. Rao and M. J. Wood (2013). "RISC in PD: the impact of microRNAs in Parkinson's disease cellular and molecular pathogenesis." Front Mol Neurosci **6**: 40.

Horvath, P. and R. Barrangou (2010). "CRISPR/Cas, the immune system of bacteria and archaea." Science **327**(5962): 167-170.

Hsu, P. D., E. S. Lander and F. Zhang (2014). "Development and applications of CRISPR-Cas9 for genome engineering." Cell **157**(6): 1262-1278.

Huang, Y., L. Cheung, D. Rowe and G. Halliday (2004). "Genetic contributions to Parkinson's disease." Brain Res Brain Res Rev **46**(1): 44-70.

Imai, Y., S. Gehrke, H.-Q. Wang, R. Takahashi, K. Hasegawa, E. Oota and B. Lu (2008). Phosphorylation of 4E-BP by LRRK2 affects the maintenance of dopaminergic neurons in Drosophila. EMBO J. **27:** 2432-2443.

Izaurralde, E., J. Stepinski, E. Darzynkiewicz and I. W. Mattaj (1992). "A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs." J Cell Biol **118**(6): 1287-1295.

Jackson, R. J., C. U. Hellen and T. V. Pestova (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol **11**(2): 113-127.

Jang, H., D. A. Boltz, R. G. Webster and R. J. Smeyne (2009). "Viral parkinsonism." Biochim Biophys Acta **1792**(7): 714-721.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science **337**(6096): 816-821.

Jinek, M., A. East, A. Cheng, S. Lin, E. Ma and J. Doudna (2013). "RNA-programmed genome editing in human cells." Elife **2**: e00471.

Kachergus, J., I. F. Mata, M. Hulihan, J. P. Taylor, S. Lincoln, J. Aasly, J. M. Gibson, O. A. Ross, T. Lynch, J. Wiley, H. Payami, J. Nutt, D. M. Maraganore, K. Czyzewski, M. Styczynska, Z. K. Wszolek, M. J. Farrer and M. Toft (2005). "Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations." Am J Hum Genet **76**(4): 672-680.

Kalinderi, K., S. Bostantjopoulou, Z. Katsarou, M. Dimikiotou and L. Fidani (2015). "D620N mutation in the VPS35 gene and R1205H mutation in the EIF4G1 gene are uncommon in the Greek population." Neurosci Lett **606**: 113-116.

Karess, R. E. and G. M. Rubin (1984). "Analysis of P transposable element functions in Drosophila." Cell **38**(1): 135-146.

Kitada, T., S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura, S. Minoshima, M. Yokochi, Y. Mizuno and N. Shimizu (1998). "Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism." Nature **392**(6676): 605-608.

Klemenz, R., U. Weber and W. J. Gehring (1987). "The white gene as a marker in a new P-element vector for gene transfer in Drosophila." Nucleic Acids Res **15**(10): 3947-3959.

Koushika, S. P., M. J. Lisbin and K. White (1996). "ELAV, a Drosophila neuron-specific protein, mediates the generation of an alternatively spliced neural protein isoform." Curr Biol **6**(12): 1634-1641.

Kruger, R., W. Kuhn, T. Muller, D. Woitalla, M. Graeber, S. Kosel, H. Przuntek, J. T. Epplen, L. Schols and O. Riess (1998). "Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease." Nat Genet **18**(2): 106-108.

Lang, A. E. and A. M. Lozano (1998). "Parkinson's disease. First of two parts." N Engl J Med **339**(15): 1044-1053.

Lang, A. E. and A. M. Lozano (1998). "Parkinson's disease. Second of two parts." N Engl J Med **339**(16): 1130-1143.

Lavara-Culebras, E. and N. Paricio (2007). "Drosophila DJ-1 mutants are sensitive to oxidative stress and show reduced lifespan and motor deficits." Gene **400**(1-2): 158-165.

Lee, B. D., J. H. Shin, J. VanKampen, L. Petrucelli, A. B. West, H. S. Ko, Y. I. Lee, K. A. Maguire-Zeiss, W. J. Bowers, H. J. Federoff, V. L. Dawson and T. M. Dawson (2010). "Inhibitors of leucine-rich repeat kinase-2 protect against models of Parkinson's disease." Nat Med **16**(9): 998-1000.

Lesage, S., C. Condroyer, S. Klebe, E. Lohmann, F. Durif, P. Damier, F. Tison, M. Anheim, A. Honore, F. Viallet, A. M. Bonnet, A. M. Ouvrard-Hernandez, M. Vidailhet, A. Durr and A. Brice (2012). "EIF4G1 in familial Parkinson's disease: pathogenic mutations or rare benign variants?" Neurobiol Aging **33**(9): 2233.e2231-2233.e2235.

Li, K., B. S. Tang, J. F. Guo, M. X. Lou, Z. Y. Lv, Z. H. Liu, Y. Tian, C. Y. Song, K. Xia and X. X. Yan (2013). "Analysis of EIF4G1 in ethnic Chinese." BMC Neurol **13**: 38.

Lieber, M. R. (2010). "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway." Annu Rev Biochem **79**: 181-211.

Liew, C. W., A. Assmann, A. T. Templin, J. C. Raum, K. L. Lipson, S. Rajan, G. Qiang, J. Hu, D. Kawamori, I. Lindberg, L. H. Philipson, N. Sonenberg, A. B. Goldfine, D. A. Stoffers, R. G. Mirmira, F. Urano and R. N. Kulkarni (2014). "Insulin regulates carboxypeptidase E by modulating translation initiation scaffolding protein eIF4G1 in pancreatic beta cells." Proc Natl Acad Sci U S A **111**(22): E2319-2328.

Liu, Z., X. Wang, Y. Yu, X. Li, T. Wang, H. Jiang, Q. Ren, Y. Jiao, A. Sawa, T. Moran, C. A. Ross, C. Montell and W. W. Smith (2008). "A Drosophila model for LRRK2-linked parkinsonism." Proc Natl Acad Sci U S A **105**(7): 2693-2698.

Lomakin, I. B., C. U. Hellen and T. V. Pestova (2000). "Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation." Mol Cell Biol **20**(16): 6019-6029.

Ma, T., M. A. Trinh, A. J. Wexler, C. Bourbon, E. Gatti, P. Pierre, D. R. Cavener and E. Klann (2013). Suppression of eIF2α kinases alleviates Alzheimer's disease–related plasticity and memory deficits. Nature Neuroscience.

Magadan, A. H., M. E. Dupuis, M. Villion and S. Moineau (2012). "Cleavage of phage DNA by the Streptococcus thermophilus CRISPR3-Cas system." PLoS One **7**(7): e40913.

Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost and E. V. Koonin (2011). "Evolution and classification of the CRISPR-Cas systems." Nat Rev Microbiol **9**(6): 467-477.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville and G. M. Church (2013). "RNA-guided human genome engineering via Cas9." Science **339**(6121): 823-826.

Marchione, R., S. A. Leibovitch and J. L. Lenormand (2013). "The translational factor eIF3f: the ambivalent eIF3 subunit." Cell Mol Life Sci **70**(19): 3603-3616.

Marsh, J. L. and L. M. Thompson (2006). "Drosophila in the study of neurodegenerative disease." Neuron **52**(1): 169-178.

Martin, I., V. L. Dawson and T. M. Dawson (2011). "Recent advances in the genetics of Parkinson's disease." Annu Rev Genomics Hum Genet **12**: 301-325.

Martin, I., J. W. Kim, B. D. Lee, H. C. Kang, J. C. Xu, H. Jia, J. Stankowski, M. S. Kim, J. Zhong, M. Kumar, S. A. Andrabi, Y. Xiong, D. W. Dickson, Z. K. Wszolek, A. Pandey, T. M. Dawson and V. L. Dawson (2014). "Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease." Cell **157**(2): 472-485.

Maruyama, T., S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram and H. L. Ploegh (2015). "Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining." Nat Biotechnol **33**(5): 538-542.

Mata, I. F., J. P. Taylor, J. Kachergus, M. Hulihan, C. Huerta, C. Lahoz, M. Blazquez, L. M. Guisasola, C. Salvador, R. Ribacoba, C. Martinez, M. Farrer and V. Alvarez (2005). "LRRK2 R1441G in Spanish patients with Parkinson's disease." Neurosci Lett **382**(3): 309-311.

Moore, D. J., A. B. West, V. L. Dawson and T. M. Dawson (2005). "Molecular pathophysiology of Parkinson's disease." Annu Rev Neurosci **28**: 57-87.

Moreno, J. A., H. Radford, D. Peretti, J. R. Steinert, N. Verity, M. G. Martin, M. Halliday, J. Morgan, D. Dinsdale, C. A. Ortori, D. A. Barrett, P. Tsaytler, A. Bertolotti, A. E. Willis, M. Bushell and G. R. Mallucci (2012). Sustained translational repression by eIF2α-P mediates prion neurodegeneration. Nature, Nature Publishing Group. **485:** 507-511.

Nichols, W. C., N. Pankratz, D. Hernandez, C. Paisan-Ruiz, S. Jain, C. A. Halter, V. E. Michaels, T. Reed, A. Rudolph, C. W. Shults, A. Singleton and T. Foroud (2005). "Genetic screening for a single common LRRK2 mutation in familial Parkinson's disease." Lancet **365**(9457): 410-412.

Nishioka, K., M. Funayama, C. Vilarino-Guell, K. Ogaki, Y. Li, R. Sasaki, Y. Kokubo, S. Kuzuhara, J. M. Kachergus, S. A. Cobb, H. Takahashi, Y. Mizuno, M. J. Farrer, O. A. Ross and N. Hattori (2014). "EIF4G1 gene mutations are not a common cause of Parkinson's disease in the Japanese population." Parkinsonism Relat Disord **20**(6): 659-661.

Nuytemans, K., G. Bademci, V. Inchausti, A. Dressen, D. D. Kinnamon, A. Mehta, L. Wang, S. Zuchner, G. W. Beecham, E. R. Martin, W. K. Scott and J. M. Vance (2013). "Whole exome sequencing of rare variants in EIF4G1 and VPS35 in Parkinson disease." Neurology **80**(11): 982-989.

Obeso, J. A., M. C. Rodriguez-Oroz, B. Benitez-Temino, F. J. Blesa, J. Guridi, C. Marin and M. Rodriguez (2008). "Functional organization of the basal ganglia: therapeutic implications for Parkinson's disease." Mov Disord **23 Suppl 3**: S548-559.

Paquet, D., D. Kwart, A. Chen, A. Sproul, S. Jacob, S. Teo, K. M. Olsen, A. Gregg, S. Noggle and M. Tessier-Lavigne (2016). "Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9." Nature **533**(7601): 125-129.

Parkinson, J. (2002). "An essay on the shaking palsy. 1817." J Neuropsychiatry Clin Neurosci **14**(2): 223-236; discussion 222.

Pelletier, J. and N. Sonenberg (1988). "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA." Nature **334**(6180): 320-325.

Perez, E. E., J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, Y. L. Lee, D. Y. Guschin, I. Rupniewski, A. J. Waite, C. Carpenito, R. G. Carroll, J. S. Orange, F. D. Urnov, E. J. Rebar, D. Ando, P. D. Gregory, J. L. Riley, M. C. Holmes and C. H. June (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases." Nat Biotechnol **26**(7): 808-816.

Pfeiffer, B. D., T. T. Ngo, K. L. Hibbard, C. Murphy, A. Jenett, J. W. Truman and G. M. Rubin (2010). "Refinement of tools for targeted gene expression in Drosophila." Genetics **186**(2): 735-755.

Platt, R. J., S. Chen, Y. Zhou, M. J. Yim, L. Swiech, H. R. Kempton, J. E. Dahlman, O. Parnas, T. M. Eisenhaure, M. Jovanovic, D. B. Graham, S. Jhunjhunwala, M. Heidenreich, R. J. Xavier, R. Langer, D. G. Anderson, N. Hacohen, A. Regev, G. Feng, P. A. Sharp and F. Zhang (2014). "CRISPR-Cas9 knockin mice for genome editing and cancer modeling." Cell **159**(2): 440-455.

Polymeropoulos, M. H., C. Lavedan, E. Leroy, S. E. Ide, A. Dehejia, A. Dutra, B. Pike, H. Root, J. Rubenstein, R. Boyer, E. S. Stenroos, S. Chandrasekharappa, A. Athanassiadou, T. Papapetropoulos, W. G. Johnson, A. M. Lazzarini, R. C. Duvoisin, G. Di Iorio, L. I. Golbe and R. L. Nussbaum (1997). "Mutation in the alpha-synuclein gene identified in families with Parkinson's disease." Science **276**(5321): 2045-2047.

Ramirez-Valle, F., S. Braunstein, J. Zavadil, S. C. Formenti and R. J. Schneider (2008). "eIF4GI links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy." J Cell Biol **181**(2): 293-307.

Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott and F. Zhang (2013). "Genome engineering using the CRISPR-Cas9 system." Nat Protoc **8**(11): 2281-2308.

Recasens, A. and B. Dehay (2014). "Alpha-synuclein spreading in Parkinson's disease." Front Neuroanat **8**: 159.

Riemensperger, T., G. Isabel, H. Coulom, K. Neuser, L. Seugnet, K. Kume, M. Iche-Torres, M. Cassar, R. Strauss, T. Preat, J. Hirsh and S. Birman (2011). "Behavioral consequences of dopamine deficiency in the Drosophila central nervous system." Proc Natl Acad Sci U S A **108**(2): 834-839.

Rogers, A. N., D. Chen, G. McColl, G. Czerwieniec, K. Felkey, B. W. Gibson, A. Hubbard, S. Melov, G. J. Lithgow and P. Kapahi (2011). "Life span extension via eIF4G inhibition is mediated by posttranscriptional remodeling of stress response gene expression in C. elegans." Cell Metab **14**(1): 55-66.

Rubin, G. M. and A. C. Spradling (1982). "Genetic transformation of Drosophila with transposable element vectors." Science **218**(4570): 348-353.

Santini, E., T. N. Huynh, A. F. MacAskill, A. G. Carter, P. Pierre, D. Ruggero, H. Kaphzan and E. Klann (2013). Exaggerated translation causes synaptic and behavioural aberrations associated with autism. Nature, Nature Publishing Group. **493:** 411-415.

Sapranauskas, R., G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath and V. Siksnys (2011). "The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli." Nucleic Acids Res **39**(21): 9275-9282.

Schulte, E. C., B. Mollenhauer, A. Zimprich, B. Bereznai, P. Lichtner, D. Haubenberger, W. Pirker, T. Brucke, M. J. Molnar, A. Peters, C. Gieger, C. Trenkwalder and J. Winkelmann (2012). "Variants in eukaryotic translation initiation factor 4G1 in sporadic Parkinson's disease." Neurogenetics **13**(3): 281-285.

Sharma, A., C. A. Hoeffer, Y. Takayasu, T. Miyawaki, S. M. McBride, E. Klann and R. S. Zukin (2010). "Dysregulation of mTOR signaling in fragile X syndrome." J Neurosci **30**(2): 694-702.

Shin, J. H., H. S. Ko, H. Kang, Y. Lee, Y. I. Lee, O. Pletinkova, J. C. Troconso, V. L. Dawson and T. M. Dawson (2011). "PARIS (ZNF746) repression of PGC-1alpha contributes to neurodegeneration in Parkinson's disease." Cell **144**(5): 689-702.

Silvera, D., R. Arju, F. Darvishian, P. H. Levine, L. Zolfaghari, J. Goldberg, T. Hochman, S. C. Formenti and R. J. Schneider (2009). "Essential role for eIF4GI overexpression in the pathogenesis of inflammatory breast cancer." Nat Cell Biol **11**(7): 903-908.

Singleton, A. B., M. Farrer, J. Johnson, A. Singleton, S. Hague, J. Kachergus, M. Hulihan, T. Peuralinna, A. Dutra, R. Nussbaum, S. Lincoln, A. Crawley, M. Hanson, D. Maraganore, C. Adler, M. R. Cookson, M. Muenter, M. Baptista, D. Miller, J. Blancato, J. Hardy and K. Gwinn-Hardy (2003). "alpha-Synuclein locus triplication causes Parkinson's disease." Science **302**(5646): 841.

Smith, W. W., Z. Pei, H. Jiang, V. L. Dawson, T. M. Dawson and C. A. Ross (2006). Kinase activity of mutant LRRK2 mediates neuronal toxicity. Nature Neuroscience. **9:** 1231-1233.

Sonenberg, N. and A. G. Hinnebusch (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell, Elsevier Inc. **136:** 731-745.

Sonenberg, N., M. A. Morgan, W. C. Merrick and A. J. Shatkin (1978). "A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA." Proc Natl Acad Sci U S A **75**(10): 4843-4847.

Spillantini, M. G., R. A. Crowther, R. Jakes, N. J. Cairns, P. L. Lantos and M. Goedert (1998). "Filamentous alpha-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies." Neurosci Lett **251**(3): 205-208.

Spillantini, M. G., R. A. Crowther, R. Jakes, M. Hasegawa and M. Goedert (1998). "alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies." Proc Natl Acad Sci U S A **95**(11): 6469-6473.

Spillantini, M. G., M. L. Schmidt, V. M. Lee, J. Q. Trojanowski, R. Jakes and M. Goedert (1997). "Alpha-synuclein in Lewy bodies." Nature **388**(6645): 839-840.

Spira, P. J., D. M. Sharpe, G. Halliday, J. Cavanagh and G. A. Nicholson (2001). "Clinical and pathological features of a Parkinsonian syndrome in a family with an Ala53Thr alpha-synuclein mutation." Ann Neurol **49**(3): 313-319.

Strongosky, A. J., M. Farrer and Z. K. Wszolek (2008). "Are Parkinson disease patients protected from some but not all cancers?" Neurology **71**(20): 1650; author reply 1650-1651.

Sudhaman, S., M. Behari, S. T. Govindappa, U. B. Muthane, R. C. Juyal and B. K. Thelma (2013). "VPS35 and EIF4G1 mutations are rare in Parkinson's disease among Indians." Neurobiol Aging **34**(10): 2442.e2441-2443.

Svitkin, Y. V., B. Herdy, M. Costa-Mattioli, A. C. Gingras, B. Raught and N. Sonenberg (2005). "Eukaryotic translation initiation factor 4E availability controls the switch between cap-dependent and internal ribosomal entry site-mediated translation." Mol Cell Biol **25**(23): 10556-10565.

Tarun, S. Z., Jr. and A. B. Sachs (1996). "Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G." Embo j **15**(24): 7168-7177.

Thomas, K. R., K. R. Folger and M. R. Capecchi (1986). "High frequency targeting of genes to specific sites in the mammalian genome." Cell **44**(3): 419-428.

Trinh, J., I. Guella and M. J. Farrer (2014). "Disease penetrance of late-onset parkinsonism: a meta-analysis." JAMA Neurol **71**(12): 1535-1539.

Tu, L., Z. Liu, X. He, Y. He, H. Yang, Q. Jiang, S. Xie, G. Xiao, X. Li, K. Yao and W. Fang (2010). "Over-expression of eukaryotic translation initiation factor 4 gamma 1 correlates with tumor progression and poor prognosis in nasopharyngeal carcinoma." Mol Cancer **9**: 78.

Tucci, A., G. Charlesworth, U. M. Sheerin, V. Plagnol, N. W. Wood and J. Hardy (2012). "Study of the genetic variability in a Parkinson's Disease gene: EIF4G1." Neurosci Lett **518**(1): 19-22.

Vanhauwaert, R. and P. Verstreken (2015). "Flies with Parkinson's disease." Exp Neurol **274**(Pt A): 42-51.

Venken, K. J. and H. J. Bellen (2007). "Transgenesis upgrades for Drosophila melanogaster." Development **134**(20): 3571-3584.

Villa, N., A. Do, J. W. Hershey and C. S. Fraser (2013). "Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome." J Biol Chem **288**(46): 32932-32940.

Vosler, P. S., Y. Gao, C. S. Brennan, A. Yanagiya, Y. Gan, G. Cao, F. Zhang, S. J. Morley, N. Sonenberg, M. V. Bennett and J. Chen (2011). "Ischemia-induced calpain activation causes eukaryotic (translation) initiation factor 4G1 (eIF4GI) degradation, protein synthesis inhibition, and neuronal death." Proc Natl Acad Sci U S A **108**(44): 18102-18107.

Vosshall, L. B., A. M. Wong and R. Axel (2000). "An olfactory sensory map in the fly brain." Cell **102**(2): 147-159.

Wakabayashi, K., K. Tanji, F. Mori and H. Takahashi (2007). "The Lewy body in Parkinson's disease: molecules implicated in the formation and degradation of alpha-synuclein aggregates." Neuropathology **27**(5): 494-506.

Wang, H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang and R. Jaenisch (2013). "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering." Cell **153**(4): 910-918.

West, A. B., V. L. Dawson and T. M. Dawson (2005). "To die or grow: Parkinson's disease and cancer." Trends Neurosci **28**(7): 348-352.

Yan, R. and R. E. Rhoads (1995). "Human protein synthesis initiation factor eIF-4 gamma is encoded by a single gene (EIF4G) that maps to chromosome 3q27-qter." Genomics **26**(2): 394-398.

Yang, H., H. Wang, C. S. Shivalila, A. W. Cheng, L. Shi and R. Jaenisch (2013). "One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering." Cell **154**(6): 1370-1379.

Yu, C., Y. Liu, T. Ma, K. Liu, S. Xu, Y. Zhang, H. Liu, M. La Russa, M. Xie, S. Ding and L. S. Qi (2015). "Small molecules enhance CRISPR genome editing in pluripotent stem cells." Cell Stem Cell **16**(2): 142-147.

Zarranz, J. J., J. Alegre, J. C. Gomez-Esteban, E. Lezcano, R. Ros, I. Ampuero, L. Vidal, J. Hoenicka, O. Rodriguez, B. Atares, V. Llorens, E. Gomez Tortosa, T. del Ser, D. G. Munoz and J. G. de Yebenes (2004). "The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia." Ann Neurol **55**(2): 164-173.

Zhao, Y., P. Ho, K. M. Prakash, J. N. Foo, J. J. Liu, W. L. Au, L. C. Tan and E. K. Tan (2013). "Analysis of EIF4G1 in Parkinson's disease among Asians." Neurobiol Aging **34**(4): 1311.e1315-1316.

Curriculum Vitae

Hao Jia **January 27, 2017**

Educational History

Ph.D. expected 2017 Cellular and Molecular Physiology Johns Hopkins University

Mentors: Valina Dawson, Ph.D. and Ted Dawson, M.D. Ph.D.

Thesis project: The Role of Parkinson's Disease Associated Mutations in eIF4G1 on Protein translation and Neurodegeneration

B.S. 2011 Biological Science Nankai University

Other Professional Experiences

Research rotation: Mar. 2012-Jun. 2012 Lab of Roger Reeves, Ph.D.

McKusick-Nathans Institute for Genetic, Johns Hopkins School of Medicine

Research rotation: Sep. 2011-Dec. 2011 Lab of Wong, Guang William, Ph.D.

Center for Metabolism & Obesity Research, Johns Hopkins School of Medicine

Research assistant: 2008-2011 Lab of Zhinan Yin, M.D., Ph.D.

Laboratory of Immunoregulation Research, College of Life Sciences, Nankai University

Publications

Ruo-Jing Li, Jing Xu, Chenglai Fu, Jing Zhang, Yujun George Zheng, **Hao Jia**, Jun O Liu. Regulation of mTORC1 by lysosomal calcium and calmodulin. ***Elife.*** 2016 Oct 27;5.

Ian Martin, Jungwoo Wren Kim, Byoung Dae Lee, Hochul Kang, Jin-Chong Xu, **Hao Jia**, Jeannette Stankowski, Min-Sik Kim, Jun Zhong, Manoj Kumar, Shaida Andrabi, Dennis W. Dickson, Zbigniew K. Wszolek, Akhilesh Pandey, Ted M. Dawson, and Valina

L. Dawson Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease. ***Cell***. 2014 Apr 10;157(2):472-85.

Jianlei Hao, Siyuan Dong, Siyuan Xia, Weifeng He, **Hao Jia**, Song Zhang, Jun Wei, Rebecca L. O'Brien, Willi K. Born, Zhenzhou Wu, Puyue Wang, Jihong Han, Zhangyong Hong, Liqing Zhao and Zhinan Yin. Regulatory Role of Vγ1 γδ T Cells in Tumor Immunity through IL-4 Production. ***J Immunol.*** 2011 Nov 15;187(10):4979-86.

Awards and Honors

Second Prize of Excellent Undergraduate Scholarship of Nankai University, 2007-2008.

Third Prize of Excellent Undergraduate Scholarship of Nankai University, 2008-2009.

Second Prize of Excellent Undergraduate Scholarship of Nankai University, 2009-2010.

Excellent Student Leader of Nankai University, 2007-2008.

Triple-A" Outstanding Student (with a good morality, a good intelligence and a good health) of Nankai University, 2009-2010.

Service and Leadership

Sep. 2012	Tutor for course *Biophysical and Biochemical Principles*
Feb.-Apr. 2014	Teaching assistant for course *Pathways and Regulation*s
2014-2015	Director of Case Competition of Johns Hopkins Graduate Consulting Club
2015-2016	Director of Operations of Johns Hopkins Graduate Consulting Club